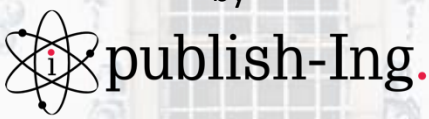
CPSL

Proceedings of the
Conference on Production Systems and Logistics

CPSL 2026

14th – 17th April 2026
Instituto Superior de Engenharia do Porto (ISEP)
Porto, Portugal

by



www.cpsl-conference.com

CPSL

Conference on Production Systems and Logistics

Instituto Superior de Engenharia do Porto (ISEP)
Porto, Portugal

14th – 17th April 2026
Proceedings

Editors:
Dr.-Ing. David Herberger
Marco Hübner, M. Sc.

https://doi.org/10.15488/20968
ISSN: 2701-6277

publish-Ing.

© Institute for Production and Logistics Research GbR
Herberger, Hübner & Beuse

Senator-Burda-Straße 31a,
77654 Offenburg
Germany

E-mail: info@cpsl-conference.com
USt-IdNr.: DE324911115

Conference Website www.cpsl-conference.com

All papers of the proceedings are made available to the public under the following Creative Commons license for an unlimited period of time:
Creative Commons Attribution 3.0 Germany

Link to summary and binding version of the license text:
http://creativecommons.org/licenses/by/3.0/de/

If contents of the proceedings are used for further work, these are to be clearly referenced in accordance with good scientific practice. The recommended citation is:

'Name 1, Given Name 1'; 'Name 2, Given Name 2';'…': 'Titel'. In: Herberger, D.; Hübner, M. (Eds.): Proceedings of the Conference on Production Systems and Logistics (CPSL 2026), 2026, publish-Ing., pp. XX-XX.

Foreword

We are pleased to present the proceedings of the eighth Conference on Production Systems and Logistics (CPSL 2026), hosted at the Instituto Superior de Engenharia do Porto (ISEP) in Porto, Portugal.

Porto, a city renowned for its rich heritage and dynamic economic landscape, offers an inspiring environment for academic exchange and innovation. As one of Portugal's key industrial and technological hubs, Porto combines a strong tradition in engineering with a forward-looking approach to research and development. Participants of CPSL 2026 are invited to engage with the latest advancements in production systems and logistics while experiencing a city that seamlessly blends historic character with modern infrastructure. The collaborative spirit of Porto, together with the academic excellence of ISEP, provides an ideal setting to foster meaningful dialogue and international cooperation.

CPSL 2026 continues to build on a strong and growing network of academic and industry partners. We are particularly grateful for the ongoing commitment of our conference committee members and partner institutions, whose continuous support has been instrumental in shaping CPSL into a recognized platform for high-quality research exchange. With each edition, we further strengthen this network and expand opportunities for collaboration among participants.

The conference features a diverse program of paper presentations and discussions, offering insights into current research trends and emerging challenges in production and logistics. As in previous years, all accepted contributions are included in these proceedings, reflecting the breadth and depth of research within our international community.

CPSL remains dedicated to providing a platform for researchers and practitioners to connect, exchange ideas, and explore innovative solutions to contemporary challenges. The formal sessions, informal discussions, and networking opportunities throughout the conference are central to advancing knowledge and fostering long-term collaboration in the field.

We would like to express our sincere appreciation to all those who contributed to the success of CPSL 2026. We thank our reviewers for their thorough and constructive evaluations, our keynote speakers for their inspiring contributions, and the entire team at ISEP for their outstanding support in hosting this event. Finally, we extend our gratitude to all participants for their valuable contributions and engagement.

We hope that CPSL 2026 in Porto will be a rewarding and memorable experience for all, and we look forward to continuing this journey together in future editions of the conference.

Dr.-Ing. David Herberger
Conference Chair

M. Sc. Marco Hübner
Conference Chair

Review Process

The Conference on Production Systems and Logistics (CPSL) is an international forum for the scientific exchange on current findings in the field of production engineering.

For the submission of a paper, an abstract had to be uploaded considering the following main topics:

- Automation
- Business Administration
- Digitalization & Industry 4.0
- Factory Planning
- Knowledge and Change Management
- Lean & Operational Excellence
- Machine Learning & Data Mining
- Production Planning & Control
- Process Management
- Supply Chain Management
- Sustainability
- Technology Driver

The submitted abstracts were evaluated in an internal review process, whereby successful submissions were invited to upload a full paper. Full papers had to adhere to a specific template and format provided on the CPSL website.

The submitted full papers were reviewed in a two-stage peer review process by experienced scientists from renowned research institutions as well as authors of other submitted papers. This process ensures a constant and high quality as well as the influence of all participants on the papers and reviews. Consequently, each paper submitted was sent to at least two reviewers, with a third reviewer being requested in case of non-consensus between the first two reviewers, if applicable.

The reviewers were asked to review the submitted papers on the basis of a provided evaluation template and were encouraged to give detailed comments and suggestions for improvement. Among others, the following key questions were considered:

- Does the title reflect the contents of the paper?
- How do you rate the comprehensibility and logical presentation of the paper?
- How do you assess the relevance and originality of the topic described?
- How do you assess the practical relevance of the paper?
- Do you consider the work a proof of a thorough research and knowledge of the latest literature in the field of research?
- Are the conclusions clear and valid?

After completion of the reviews, all authors were given sufficient time to improve their papers according to the remarks of the reviewers. All papers that received major remarks in the first review loop were reviewed again in a second review loop and were also accepted, provided that the remarks were conscientiously incorporated.
For more information on the review process and the “Publication Ethics and Publication Malpractice Statement” please visit the conference website.

CPSL

Acknowledgements

Our sincere thanks goes to our outstanding supporters who made this great event possible.

Publishing

publish-Ing.

TIB Leibniz-Informationszentrum Technik und Naturwissenschaften Universitätsbibliothek

Scientific Societies

WGP Wissenschaftliche Gesellschaft für Produktionstechnik

Venue & Co-Host

ISEP

Reviews & Scientific Support

Texas A&M University Texarkana

fir an der RWTH Aachen

Fraunhofer IAO

IFA Institut für Fabrikanlagen und Logistik

UBC Sauder School of Business

udla

Universidad Austral

Universidad de Lima

KEDGE Business School

TIB Leibniz-Informationszentrum Technik und Naturwissenschaften Universitätsbibliothek

University of Exeter

LPS Lehrstuhl für Produktionssysteme

Universität Stuttgart Institut für Industrielle Fertigung und Fabrikbetrieb IFF

NUS National University of Singapore

University of Birmingham

University of Twente.

ISM Institute for Sustainable Manufacturing

Reutlingen University Nachhaltigkeit und Technologie NXT

Institut für Technologiemanagement Universität St.Gallen

FAPS

Tecnológico de Monterrey

Universität Kassel

Curtin University

WZL RWTH Aachen University

Universität Augsburg University

pfp

TU Darmstadt

Technical University of Leoben

UM Universidad de Montevideo

LFO

WGP Wissenschaftliche Gesellschaft für Produktionstechnik

Universidade Federal de Santa Catarina

Keynote Speakers 2026

A special THANK YOU goes to our outstanding Keynote speakers, who joined the conference and inspired our participants beyond their presentations with cutting-edge and highly interesting topics

Prof. Dr. Paul Molenda
Hof University, GER

Next Level Value Creation: How AI and Digital Innovation Transform Products, Processes and Business

Digital transformation shifts industrial value creation from linear chains to intelligent and iterative networks. This keynote highlights how AI-enabled products, planning systems and flexible production concepts open up new forms of value creation and outlines the organizational capabilities needed to strategically unlock these potentials.

Prof. Dr. Lourdes Perea Muñoz
Universidad Austral, ARG

Beyond Employability: Educating Professionals for Lifelong Relevance in the Age of AI and Industrial Change

In a landscape where digital technologies and AI are transforming industry and the way value is created, the way we work—and the professional competencies required—are evolving as well. This keynote reflects on the changing role of universities, the meaning of work, and the curricular adaptations needed to prepare professionals who can remain relevant and contribute meaningfully in an environment defined by accelerated change and the rapid obsolescence of technical knowledge.

Prof. Dr. Paulo Ávila
Instituto Superior De Engenharia Do Porto, POR

Why and how to boost agile and virtual enterprises to attend over capacity production systems: Insights and Strategies

Overcapacity in production systems is traditionally viewed as a structural inefficiency, associated with idle resources, financial pressure, and reduced margins. However, when strategically managed through agile and virtual enterprise models, overcapacity can be transformed from a cost burden into a powerful market enabler. Rather than representing waste, surplus capacity can become a source of responsiveness, innovation, and competitive advantage. To unlock these benefits, the keynote highlights key enablers such as: efficient resource selection, digital integration, real-time data sharing, trust-based and decision governance models.

Table of Contents

1	Assessing the Environmental Benefits of Remanufactured Proton Exchange Membrane Fuel Cell Stacks *Marie Beller, Manuel Querbach, Frank Döpper*	1
2	Towards Explainable Knowledge And Behavior In Social Robotics Using Fuzzy Techniques: A Systematic Literature Review *Raoul Zebisch, Johannes Schilp*	16
3	Implementation and Validation of Elastic Transformation Engine for Asset Administration Shells *Benjamin Goetz, Matthias Schneider, Daniel Schel, Thomas Bauernhansl*	32
4	Self-Description Of Heterogeneous System Landscapes In Manufacturing Environments *Matthias Schneider, Benjamin Götz, Daniel Schel, Thomas Bauernhansl*	43
5	Future Competitiveness In Industrial Wood Processing: A Survey-Based Analysis Of Economic, Technological and Ecological Developments *D. Roppelt, J. Schmidt, B. Thorenz, .F. Döpper*	54
6	Towards Predictive Quality Control In Large-Scale Battery Cell Finalization: A Pillar-Based Framework *Julian Wonneberger, Patric Keister, Jonathan Krauß, Simon Lux, Louis Blumenschein*	65
7	Intelligent Process Control In Paperboard Compression-Drawing Using Digital Twins And Machine Learning *Muhammad Faisal Yaqoob, Klara Liesegang, Lena Berthold, Stephan Kronenberger, Christer-Clifford Schenke, Steffen Ihlenfeldt*	76
8	Smart Factory Framework for Real-Time Inventory Control and Process Optimization *Tom Wanyama*	87
9	Decoupling Mode And Effects Analysis: Enhancing Manufacturing Resilience Through Systematic Risk Identification *Florian Pleier, Simon Rekers, Andreas Röder, Daniel Beck, Ben Lüskow, Kirsten Tracht, Johannes Schilp*	95
10	A System Readiness Assessment Of LCA Applications Within The Novel Framework Of Green Digital Twins *Gerald Feichtinger, Wolfgang Posch*	105
11	A Framework For The Knowledge-Based Decision Support System For Deviation Management On The Shop Floor *Yuxi Wang, Joachim Metternich*	117
12	LLM-Based Multimodal Prompting For Adaptive Robot Control In Production Systems *Dominik Koch, Jakob Wolber, Zhuo Shi, Bo Cheng Ji, Lucas Bretz, Alexander Geiser, Felix Baer, Martin Benfer, Florian Stamer, Gisela Lanza*	129
13	A Structured Literature Review and Interview study of human-centred Challenges in Artificial Intelligence Implementation *Verena Dondorf, Kitrujan Kokilanathan, Johanna P. Schenk, Daniela Vogel, Roman Dumitrescu*	142
14	Wear Volume Measurement Of Cutting Tools: Structured Literature Review *Benjamin Thorenz, Johannes Rauch, Frank Döpper*	155
15	Development of a Digital Model for an Educational Fischertechnik Production System Using PLC-Based Control and Tecnomatix Plant Simulation *Paul Molenda, Ben Becher, Jonathan Milz, Vihasith Ganga*	167

Table of Contents

16 Development And Evaluation Of An AI-Enhanced Smart Glove For Scannerless Order Picking In Intralogistics 177
Jan Schuhmacher, Vera Hummel

17 Artificial Intelligence-Driven Additive Manufacturing Use Cases: A Framework For Grounded Decisions 187
Konrad Schneidenbach, Carsten Feldmann

18 A Conceptual Framework for Evaluating the Transferability of Production Management Decisions to AI 204
Joshua Kaiser, Carl René Sauer, Norman Müller, Luisa Stracke, Peter Burggräf

19 Structuring Design Problems of Reconfigurable Robotic Assembly Systems 216
Elias Huber, Marc Ungen

20 Developing a Decision-Support Tool for Parameter Configuration in Automated Sorting System 226
Davide Morchia, Laura Monferdini, Daniele Gibertini, Christian Donati, Xhulia Sina, Marco Gotelli, Eleonora Bottani

21 Bridging the Synchrony Gap: A Deterministic GSMDP Approach for Integrating Deep Reinforcement Learning with Commercial Discrete Event Simulators 237
Merve Demir Varici, Hasan Tercan, Tobias Meisen

22 Logistics Model To Reduce Inventory Discrepancies By Applying Digital POKA-YOKE, RFID And WMS In The Warehouses of White-Goods Company 246
Carlos Ali Irigoin-Sánchez, Diego Vincenzo Mascaro-Levano, Jorge Antonio Corzo-Chavez, Gino Viacava-Campos

23 Toward A Coherent Definition Of The Unified Namespace In Industry 4.0 259
Janis Simons, Fabian Igelmund, Luke Hagenow, Jonas Reck, Niklas Radbruch, Tobias Schröer, Paul Katsioupis

24 Cross-Country Analysis of Digitalization of Services of General Interest: A Systematic Literature Review 270
Eva Paulina Ziehm, Nick Große, Jonas Eichholz, Leon Jungh, Marius Klinsmer Tchapga Leumen

25 Implicit Flexibility in Agent Behavior within Cloud Manufacturing 284
Robin Huwa, Marcus Albrecht, Andreas Leinenbach, Johannes Schilp

26 Reinforcement Learning for Intercropping Sequences: Transferable Decision Policies for Circular Industrial Production Systems 293
Stefan Sarawinsky, Kai Schaller, Kajan Kandiah

27 A Digital Framework For Structured Meeting Preparation In Kaizen Projects 304
Kevin Zhao, Yuxi Wang, Joachim Metternich

28 Structuring the Shift: Designing and Applying a Smart PSS Transformation Framework with the TriProGenix Approach 314
Jonas Wirth, Dennis Unic, Thomas Bauernhansl

29 From Uncertainty to Value: The Scoreboard and Feasibility Gate for Early-Stage Financial Planning in Industrial Data Ecosystems 329
Jonas Wirth, Fabian Meldt, Felix Hoffmann, Matthias Weigold, Thomas Bauernhansl

30 Design-Framework for Product-Service Systems in Subscription-Based Business Models 352
Marius Höding, Jonas Wirth, Sarah Ekanayake, Thomas Prefi, Thomas Bauernhansl

31 Development Of A Process Model For The Selection And Development Of Service-Oriented And Data-Driven Business Models In Mechanical And Plant Engineering 365
Jonas Wirth, Anka Zoefeld, Guenter Bitsch, Thomas Bauernhansl

Table of Contents

32	Challenges in Implementing Service-Oriented and Data-Driven Business Models in Mechanical and Plant Engineering: A Literature Review and Empirical Analysis *Jonas Wirth, Tony Scheible, Yasin Schneider, Thomas Bauernhansl*	376
33	From Digital Twin Towards Shop Floor: Codeless Programming Of Pick-And-Place Robots Through Extended-Reality Interaction *Andoni Rivera-Pinto, Johan Kildal*	391
34	Assessing Data Quality Dimensions And Their Impact On Production Planning and Control *Johannes Jordan, Oliver Antons, Julia Arlinghaus*	401
35	Influence Of Legacy Performance Indicators On Contemporary Production And Assembly Research *Tobias Bein, Ulf Bergmann, Oliver Antons, Julia Arlinghaus*	412
36	From Paper Production to Traction Motors: Transfer Patterns for the FOREST Digital Twin Framework *Rosario Othen, Nicolaus Klein, Florian Pohlmeyer, Christian Möbitz, Jürgen Fleischer, Thomas Gries*	426
37	Global Warming Potential Assessment for Sustainable Procurement in Production Supply Chains *Kunal Chaudhari, Oussama Etlijani, Clemens Gonnermann, Aljoscha Hieronymus, Maximilian Holland, Sönke Maatsch, Anja Munzke, Frida Schulze*	438
38	Towards Principles for Designing Lean Intralogistics Systems: Systematic Transfer of Lean Thinking to Reduce Waste in Intralogistics *Lena Langenfeld, Thomas Bauernhansl*	449
39	Development Of A Methodology To Determine Change Variance In Product Architecture Elements In The Context Of The Upgrade Circular Economy *Günther Schuh, Alexander Keuper, Frederike Hellwig, Kira Rimbach*	462
40	Concept For Designing Upgradeable Product Architectures In The Context of Upgrade Circular Economy *Frederike Hellwig, Alexander Keuper, Günther Schuh*	477
41	From Static To Dynamic Costing: Managing External Volatility and Evaluating Supply Chain Configurations in MTO Manufacturing *Lukas Fleisch, Lukas Budde, Thomas Friedli*	488
42	Evaluation Of The Realization Effort Of Product Upgrades In The Context Of The Upgrade Circular Economy *Timm Schulz-Isenbeck, Alexander Keuper, Frederike Hellwig, Günther Schuh*	498
43	Location of Micro Fulfillment Centers in a Quick Commerce Logistics Network with Consideration of Inventory Management *Zhu Zhu, Haoxun Chen*	510
44	Analyzing Actor Interaction Gaps for the Digital Product Passport in a Multi-Stakeholder Ecosystem *Doğan Efe, Elena Andrushchenko, Roland Jochem*	523
45	Requirements Analysis For Near-Ground Object Detection Of Autonomous Mobile Robots *Thorge Lackner, Julian Hermann, Christian Kuhn, Daniel Palm*	535
46	Selection Of Indicators For Integrating Ecological Sustainability Into Product Portfolio Management *Nikolai Kelbel, Alexander Keuper, Günther Schuh*	545
47	Procurement Target System of the Future *Theresa-Franziska Hinrichsen*	555
48	Forecasting Material Flow Composition in the German Returnable Beverage Container System *Leia Betting, Jakob Becker, Philipp Kempkes, Konrad Mielke, Markus Pauly*	565

Table of Contents

49	Digital Transparency In The Loop Between Shopfloor And Logistics Planning *Julian Hartwich-Fischer, Katharina Paschert, Anja Munzke, Mario Luber, Florian Karg, Meltem Akkaya*	579
50	Methodology For Dynamic Dimensioning Of Re-Assembly Factories *Günther Schuh, Esben Schukat, Nils Lehde genannt Kettler*	590
51	Synergies Among Urban Mining, Reverse Logistics, And Closed-Loop Supply Chains Towards Circular Economy: A Decision Support Framework *Kuruppuge Themiya, Kulatunga Asela, Luis Martino, Sucala Voicu Ion*	600
52	Enabling Human-AI Collaboration in a Modular Planning Environment *Esben Schukat, Lea Röhm, Markus Netzer, Günther Schuh*	611
53	A Chinese Postman Problem-Based Approach For Municipal Solid Waste Management *Akshat Kothyari, Shaunak Juvekar, Subhash C. Sarin*	620
54	An Experimental Analysis Of A Scaled Electromagnetic In-Space Launcher: A Multi-Stage Solenoid Accelerator For Nanosatellite Payload Deployment *Bhushan Lohar, Sai Raghava Pathuri, Sumedh Patil, John Wade, Mir Ali, Lance Outerbridge*	633
55	Implementing Lean Six Sigma Methodologies in an Industry 4.0 Manufacturing Setup: A Case Study *Shantall Cisneros Saldana, Heike Markus*	643
56	Optimizing Last-Mile Distribution in an Emerging-Economy Bottled Water Firm Through VRP and DDMRP Integration *Boris Castañeda-Rios, Veralucía Zamalloa-Peralta, Marcos Fernando Ruiz-Ruiz, Juan Carlos Quiroz-Flores*	653
57	Data-Driven Lean Warehousing And Hybrid Simulation To Improve Perfect Order Rate In Medical Supply SMEs *Carlos Daza-Llanos, Said Rojas-Lavado, Marcos Fernando Ruiz-Ruiz, Juan Carlos Quiroz-Flores*	664
58	Challenges And Opportunities In Remanufacturing: A Comprehensive Meta-Analysis *Jan Koller, Jonas Weidner, Vihasith Ganga, Frank Döpper*	678
59	Assessment Of Electric Off-Road Machinery For Industrial Applications In Uruguay *Daniel Jurburg, Marcelo Berglavaz, Justino Romano, Rodrigo Sanchez, Camila Perez, Micaela Montilla*	690
60	Implementation of a Digital Twin Prototype For The After-Sales Service Workshop Process *Afonso Silva, Paulo Ávila, João Bastos, Luís Pinto Ferreira, Hélio Castro, Alzira Mota*	702

8th Conference on Production Systems and Logistics

Assessing the Environmental Benefits of Remanufactured Proton Exchange Membrane Fuel Cell Stacks

Marie Beller[1], Manuel Querbach[1] , Frank Döpper[1,2]

[1]*University of Bayreuth, Chair Manufacturing and Remanufacturing Technology, Universitätsstraße 30, 95447 Bayreuth, Germany*
[2]*Fraunhofer IPA Project Group Process Innovation, Universitätsstraße 9, 95447 Bayreuth, Germany*

Abstract

The concept of a circular economy is increasingly recognized as an essential approach to achieving climate goals, largely due to its ability to reduce the carbon footprint of products. By extending the lifespan of goods or recovering valuable materials, circular practices help minimize resource use and emissions. While recycling is the most familiar element of this approach, a broader set of circular economy strategies (CE strategies) exists, many of which can deliver even greater reductions in environmental impacts. Among them, remanufacturing has proven particularly effective and is already firmly established in various sectors, e.g. for automotive components. In the field of fuel cells, however, remanufacturing is not yet fully developed. The necessary process chains still require further research and refinement to become viable on a larger scale, especially as the number of used fuel cell stacks is expected to grow in the coming years. This paper highlights the ecological benefits of remanufacturing proton exchange membrane fuel cell (PEMFC) stacks. Based on current technological advances and ongoing research, the projections of this paper suggest that including remanufacturing of PEMFC stack components could achieve an additional reduction of 5.08 % in carbon footprint compared to recycling PEMFC stacks only.

Keywords

Fuel Cells; Life Cycle Assessment; Remanufacturing; Recycling; Circular Economy

1. Introduction

Fuel cell electric vehicles (FCEV) can play a vital role in reducing the environmental impact of the automotive industry as their global warming potential (GWP) is substantially lower than combustion vehicles' GWP [1,2]. The most suitable technology for FCEVs are PEMFC stacks because of their quick start-up process, good cold start capability under low-temperature conditions, high power density, flexible operating range, lightweight compact system and operating temperature range (60 - 80 °C) [3]. PEMFC are of increasing interest as they cannot only be used for mobile but also stationary applications. This is reflected in the increasing research activities on cell design, material compositions and production processes [4]. Although the sales and investment figures showed only a gradual development, the current growth rates clearly indicate an upward trend [5,6]. In order to meet EU climate targets, research needs to be carried out regarding sustainable materials and CE strategies for retracting critical materials.

State of the art in CE strategies for PEMFCs is material recycling only. The focus is primarily on platinum recovery because of its value and increasing scarcity. However, in order to close the loop more efficiently, other strategies have to be considered as well. Remanufacturing is an approach that preserves the product form, thus saving more energy and resources compared to recycling. Remanufacturing of PEMFC stacks, more specifically its process chain, is an uprising research topic. The number of funded research projects

DOI: https://doi.org/10.15488/20969
ISSN: 2701-6277

within this field is increasing. Most of them focus on the (automated) disassembly of the stacks, as it is very challenging yet crucial to disassemble the stack and cells damage-free [7,8].
This paper aims to contribute to this development by pointing out the ecological potential of remanufacturing PEMFCs. The potential is quantified by the GWP, which was derived using life cycle assessment (LCA) according to ISO 14040 and ISO 14044. To be able to classify the figures, the GWP of a remanufactured stack is compared with the GWP of a newly produced stack.

2. Literature Review

Numerous authors have already conducted and published LCA studies on PEMFCs. Most of the work focuses on the production, usage in FCEVs and recycling of PEMFC stacks [9,10,1,11,2].

The production of PEMFC stacks was analyzed many times [9,12–15,11]. The work of EVANGELISTI ET AL. and GOTTARDO ET AL. should be highlighted, as the detailed process chains reported in the supporting material have also been incorporated into this paper. In the area of recycling, the focus is on platinum recovery [16]. Various processes were analyzed for this purpose: Hydrometallurgy, pyro-hydrometallurgy and alternative techniques such as selective electrochemical dissolution or bioleaching. Due to its high technology readiness level, LCA studies were conducted mainly for hydrometallurgical processes [17–19].

As already mentioned, other CE strategies are being developed in addition to material recycling. Since non-destructive dismantling is a prerequisite for many CE strategies, the necessary processes are currently being investigated: SCHMIDT has examined the manual disassembly steps and provides an overview of cell design changes that can facilitate disassembly as well as possible CE strategies [8]. OLDEWURTEL ET AL. analyzed the general structure and layout of PEMFC stacks, with a focus on the bipolar plate (BPP), the membrane electrode assembly (MEA) and the gaskets or seals, and derived potential disassembly technologies for common PEMFC stack layouts, along with the expected limitations arising from different component and material combinations and their characteristics [20]. AL ASSADI ET AL. disassembled two different PEMFC stack types manually and derived concepts for automated disassembly, considering material characteristics and resulting challenges during disassembly [7].

While recycling processes have already been intensively investigated, remanufacturing processes – testing, sorting, cleaning, repairing, and assembling – still need to be elaborated. In particular, the ongoing research project *ReStack* should be mentioned here, which deals with the diagnosis of the stack, the separation of individual cells, but also the BPP and MEA, as well as the diagnosis and repair of the BPP [21]. As the processes are still under construction, there are no corresponding LCA studies yet. This paper contributes to closing the gap by assessing the ecological potential of remanufacturing processes for PEMFC stacks.

3. Method

A comparative LCA is performed to evaluate the ecological potential of CE strategies recycling and remanufacturing for PEMFC stacks. This is done according to the specifications of the standards ISO 14040, cf. [22], and ISO 14044, cf. [23]. These standards define the phases to be completed when preparing a LCA: Defining the objective and the scope of the study (3.1), preparing the life cycle inventory (3.2), impact assessment (0) and evaluating the results (4). The following chapters are structured accordingly. The focus of the analysis is on the impact category of global warming potential, measured in kilograms of carbon dioxide equivalent (kg CO_2 eq).

As previously outlined, the remanufacturing processes for PEMFC stacks have not yet been industrialized and there is a lack of robust data. Consequently, this study can only be seen as an indication of the ecological potential and is subject to corresponding uncertainties. In addition, the overall balance of a remanufactured PEMFC stack depends on the proportion of individual components that can be remanufactured due to their quality. Various scenarios are therefore considered in the evaluation regarding this proportion.

3.1 Goal and Scope

The goal of the LCA study is to compare the CE strategies recycling and remanufacturing for the components of PEMFC stacks in general and quantify their ecological benefit compared to the production's burden. The assessment object is a PEMFC stack (in two variants, see Table 1) consisting of a sequence of cells, which are framed with end plates and current collectors, and braced with rods and nuts. The cells themselves consist of thin layers: BPP, gaskets, gas diffusion layer (micro- and macroporous layer), catalyst layer and membrane. The latter three form the MEA, a non-destructively not dismantlable unit. Figure 1 depicts the assembly schematically.

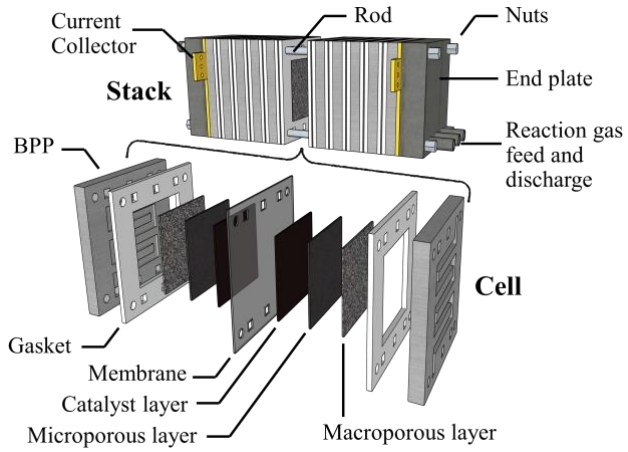


Figure 1: Structure of a PEMFC stack and the cell components

Table 1: Components in scope for two stack variants

Component	Amount (kg)		Source
	Stack 1	Stack 2	
MEA	1.9866	1.9866	[24]
Gaskets	2.3958	2.3958	[24]
Metallic BPP	11.3000	-	[24]
Graphite-Polymer BPP	-	10.8709	[24], [25]
Current collectors	0.2013	0.2013	[24]
End plates	8.1700	8.1700	[24]
Rods, Nuts	0.3267	0.3267	[24]
Total amount	24.3804	23.9513	

A 55 kW stack from the manufacturer EKPO was chosen as the reference stack, as useful details on its design have been published and LCA studies are already based on the manufacture of this stack [26,9,12,24]. The material contents in [24] were mostly reapplied. However, since industrially manufactured PEMFC stacks with both metallic and graphite composite BPP have become established on the market, the EKPO stack is modeled in two variants: with metallic and graphite composite BPP. The weight of the graphite composite BPP was adjusted according to the density and strength of graphite. The quantities can be found in Table 1. The functional unit is 1 stack.

The study covers the production (incl. raw material extraction) and end of life (EoL) phases, but not the use phase. The reason for this is that the EoL processes, which are the focus of the study, have no influence on the use phase. The production phase, which also remains unaffected, was nevertheless intentionally included to be able to classify the order of magnitude of the credits of the CE strategies. Figure 2 depicts the system boundary and gives an overview which CE strategies are technologically feasible for which components.

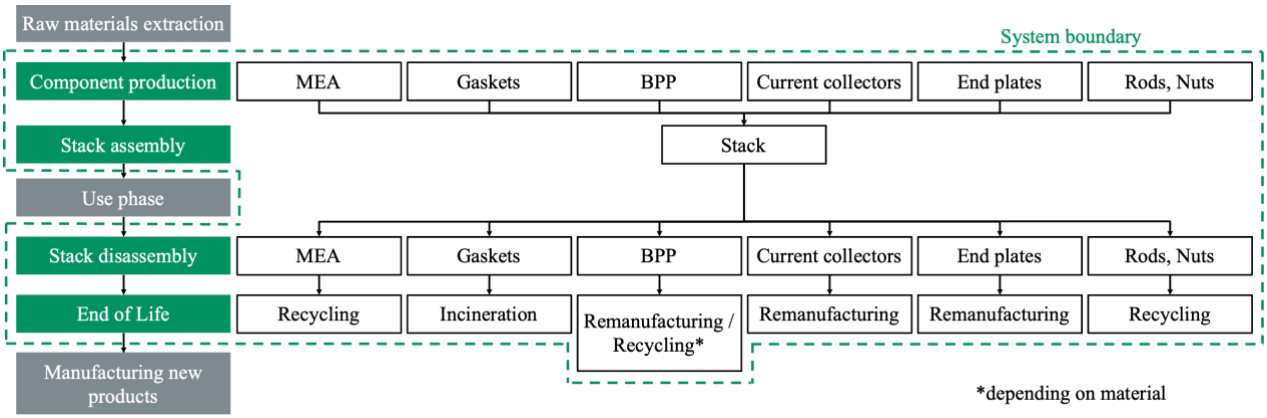


Figure 2: Processes within the system boundaries

The EoL processes of the individual PEMFC components are described briefly in the following, for the flow diagrams see Figure 4 in appendix. The MEA cannot be dismantled without damage; hence it can only be recycled. Its multi-stage recycling primarily serves to recover the platinum contained in the catalyst layer, for example by means of hydrometallurgy, cf. [27]. The gaskets are usually thermally recycled [28]. Bracing components (rods and nuts) are fed into conventional steel recycling. It is assumed that current collectors can be recycled in the same way as copper current collectors from lithium-ion batteries [29]. Alternatively, they can be treated in the classic copper recycling process using electrolysis [30]. The end plates that are suitable for reuse should be cleaned using a CO_2 snow blasting process. In the case of recycling, the steel content can be recycled using classic steel recycling methods (blast furnace or electric arc furnace line), the same applies for metallic BBP [31]. Metallic BPP cannot be reused due to their thin structure and short life time [32]. Graphite composite BPP, on the other hand, can be reused. Various processes can be used for their post-processing, cf. [33], however, CO_2 snow blasting is preferable [34]. Depending on the binder, discarded graphite composite BPP can be shredded and reprocessed as granulate (thermoplastic binder) or thermally recycled (duroplastic binder) [35].

As the disassembly of the cells of a stack is a manual process according to the current state of the art, this is not accounted for in the LCA. Similarly, the assembly of the stack is also excluded from the LCA, as no reliable data is available for this either and the stacking process of the individual components is not of great importance compared to component production.

This study uses the Allocation at the Point of Substitution (APOS) allocation method. This corresponds to the avoided burden approach, in which the environmental impacts are divided between the waste-generating product and the raw materials recovered through recycling using an allocation factor. In the APOS method, the majority of the environmental impacts of recycling are attributed to the life cycle in which the recyclable material is generated as waste. [36] As the focus of this study is on the EoL of the stack, the APOS approach is chosen to be able to map the influence of recycling and remanufacturing in the product life cycle.

The data sources used are primarily the latest possible research reports and scientific journal papers. When it comes to the disposal of PEMFC stacks, the focus in the literature is on the recovery of platinum; all other materials are generally neglected. However, it should be noted that PEMFC platinum recycling has so far only been carried out on an experimental scale. There are only a few publications on the reprocessing of PEMFC stacks, which is why the modeling of remanufacturing processes is characterized by assumptions and parallels e.g. to battery remanufacturing. The technological status can therefore be described as current but not industrially tested. The same applies to the temporal scope of the foreground data. Background data was modeled using processes of the ecoinvent database (v3.9.1), hence it is also classified as currently valid. The geographical scope is the production and recycling or remanufacturing of stacks in Europe. However, it should be noted that raw materials are also sourced outside of Europe. Accordingly, the processes are modeled as european or global.

The production processes of the components and their final assembly are described in detail in other publications as [14]. The EoL processes which lead to differing results are summarized in the following paragraph.

3.2 Life Cycle Inventory

The modeling of the production processes is largely based on the publication by GOTTARDO ET AL. and EVANGELISTI ET AL., however, the platinum content was adjusted to reflect the current state of the art of PEMFC designs [9,12]. The modeling of the MEA recycling processes is based on the publication by STROPNIK ET AL. [24]. The other recycling processes and remanufacturing processes were modeled individually according to the literature research based on the materials. The database used is ecoinvent v3.9.1. The life cycle inventory (LCI) of the production process is stated in the appendix, cf. Table 7.

The remanufacturing processes are specified in Table 2 - Table 4 per kg of input, as the input quantity for the total stack depends on the quality of the individual components, i.e. how much is suitable for

remanufacturing. Different distributions are shown later in the evaluation of the scenarios. The tables list the required materials or energy, their quantities, and the ecoinvent process used to model them.

Table 2: Life cycle inventory of the remanufacturing processes of 1 kg graphite composite BPP

Material or energy	Amount	Ecoinvent process used
Inputs		
Used graphite composite BPP	1 kg	Intermediate product, individual sub-processes listed in Table 7
Liquid CO_2	327 g	Carbon dioxide, liquid {RER}| market for | APOS, U
Compressed air	0.98 m³	Compressed air, 1000 kPa gauge {RER}| market for compressed air, 1000 kPa gauge | APOS, U
Electricity	5.7 Wh	Electricity, low voltage {Europe without Switzerland}| market group for | APOS, U
Outputs		
Remanufactured graphite composite BPP	1 kg	

Table 3: Life cycle inventory of the remanufacturing processes of 1 kg current collectors

Material or energy	Amount	Ecoinvent process used
Inputs		
Used current collectors	1 kg	Intermediate product, individual sub-processes listed in
Hydrochloric acid	0.40 kg	Hydrochloric acid, without water, in 30 % solution state {RER}| market for | APOS, U
Nitric acid	0.40 kg	Nitric acid, without water, in 50 % solution state {RER w/o RU}| market for nitric acid, without water, in 50 % solution state | APOS, U
Distilled water	0.40 kg	Water, ultrapure {RER}| market for water, ultrapure | APOS, U
Electricity	75 kJ	Electricity, medium voltage {Europe without Switzerland}| market group for | APOS, U
Outputs		
Remanufactured current collectors	1 kg	
Waste Outputs		
Wastewater	1.2 kg	Wastewater, average {Europe without Switzerland}| market for wastewater, average | APOS, U

Table 4: Life cycle inventory of the remanufacturing processes of 1 kg end plates

Material or energy	Amount	Ecoinvent process used
Inputs		
Used end plates	1 kg	Intermediate product, individual sub-processes listed in
Liquid CO_2	33.3 g	Carbon dioxide, liquid {RER}| market for | APOS, U
Compressed air	0.1 m³	Compressed air, 1000 kPa gauge {RER}| market for compressed air, 1000 kPa gauge | APOS, U
Outputs		
Remanufactured end plates	1 kg	

Table 5 lists the processes used for modelling the recycling processes for the EoL stack. For the recycled Pt/C and graphite composite proxy it is assumed that they can be reused in the production process for MEA and BPP respectively. For the sake of simplicity, only the solids contained in these two intermediate products

were specified in the LCI for, whereas in the Simapro model all solvents and energy consumption associated with their manufacturing process are taken into account.

Table 5: Life cycle inventory of the recycling processes of the 55 kW stack

Material or energy	Amount		Ecoinvent process used
	Stack 1 (metal)	Stack 2 (graphite)	
Inputs			
EoL Stack	24,380.4 g	23,951.3 g	Cf.
Hydrochloric acid	2,411.37 g	2,107.4 g	Hydrochloric acid, without water, in 30 % solution state {RER}| market for | APOS, U
Nitric acid	972.4 g	972.4 g	Nitric acid, without water, in 50 % solution state {RER w/o RU}| market for nitric acid, w/o water, in 50 % solution | APOS, U
Ammonium chloride	1,281.7 g	1,281.7 g	Ammonium chloride {GLO}| market for | APOS, U
Sodium hydroxide	267.6 g	267.6 g	Sodium hydroxide, without water, in 50 % solution state {GLO}| market for |
Filter paper	131.1 g	131.1 g	Paper, woodfree, uncoated {RER}| market for | APOS, U
Water	26,762.3 g	26,762.3 g	Tap water {Europe without Switzerland}| market for | APOS, U
Hydrogen peroxide	55,935.0 g	-	Hydrogen peroxide, without water, in 50 % solution state {GLO}| market for | APOS, U
Sodium persulfate	41.6 g	-	Sodium persulfate {GLO}| market for | APOS, U
Sodium sulfate	65.1 g	-	Sodium sulfate, anhydrite {RER}| market for | APOS, U
Sulfuric acid	40,906.0 g	-	Sulfuric acid {RER}| market for sulfuric acid | APOS, U
Zinc	15.3 g	-	Zinc {GLO}| market for | APOS, U
Iron chloride	10.8 g	-	Iron (III) chloride, without water, in 40 % solution state {GLO}| market for | APOS, U
Calcium carbonate	36,725.0 g	-	Calcium carbonate, precipitated {RER}| market for calcium carbonate, precipitated | APOS, U
Electricity	4,286.7 MJ	4,190.0 MJ	Electricity, medium voltage {Europe without Switzerland}| market group for | APOS, U
Outputs			
Recycled Pt/C	29.8 g	29.8 g	Intermediate product consisting of: Carbon black {GLO}| market for | APOS, U Platinum {GLO}| market for | APOS, U
Recycled Gold	5.5 g	-	Gold {GLO}| market for | APOS, U
Recycled stainless steel (SS316)	17,034.0 g	6,299.0 g	SS316 based on Steel, chromium steel 18/8 {GLO}| market for | APOS, U
Recycled chromium steel	348.0 g	348.0 g	Steel, chromium steel 18/8 {GLO}| market for | APOS, U
Recycled graphite composite proxy	-	9,783.8 g	Intermediate product consisting of: Graphite, battery grade {GLO}| market for | APOS, U
Recycled copper	191.2 g	191.2 g	Copper, cathode {GLO}| market for | APOS, U
Waste outputs			
Hazardous waste (GDL, solvents and solids)	12,867.2 g	13,389.3g	Hazardous waste, for incineration {Europe without Switzerland}| market for hazardous waste, for incineration | APOS, U
Solid waste	1254.3 g	-	Municipal solid waste {DE}| market for municipal solid waste | APOS, U
Wastewater	314.9 l	26.8 l	Wastewater, average {Europe without Switzerland}| market for wastewater, average | APOS, U
Thermal recovery	2,395.8 g	2,395.8 g	Waste rubber, unspecified {Europe without Switzerland}| treatment of waste rubber, unspecified, municipal incineration | APOS, U
Used filter paper	131.1 g	131.1 g	Waste paper, sorted {GLO}| market for |APOS, U

3.3 Impact Assessment

The used life cycle impact assessment method is ReCiPe 2016, which is a globally used method in academia. Therefore, comparability with other studies is facilitated. For this study, the focus is on the impact category global warming potential (GWP).

Whether parts can be remanufactured or have to be recycled depends on their condition. Since this is an individual decision and cannot be determined a priori, the impact assessment is based on scenarios. In the first comparison scenario (S1), remanufacturing is not used as an EoL scenario and all components are recycled or thermally recovered. For all other scenarios, remanufacturing is considered for the graphite composite BPP, current collectors, and end plates components, but the proportion of remanufacturable parts varies: in scenario S2, 60 % are remanufacturable, in scenario S3, 80 % are remanufacturable, and in scenario S4, 90 % are remanufacturable.

Figure 3 gives an overview of the impact categories *global warming potential*, *fine particulate matter* and *stratospheric ozone depletion*. These are characterized as level I impact categories (recommended and satisfactory) by the European union [37], all other impact categories are listed in Table 8 and 9, appendix. The impacts are depicted for the different scenarios S1-S4 and divided into the contributions for production, recycling and remanufacturing. The sum of positive and negative environmental impacts is visualized by the blue line. By incorporating remanufacturing processes, the GWP of a PEMFC stack can be reduced by another 2.40 % (stack 1) / 5.08 % (stack 2). For the impact categories fine particulate matter formation (0.59 / 1.21 %) and stratospheric ozone depletion (1.33 / 2.45 %) the effect is less noticeable.

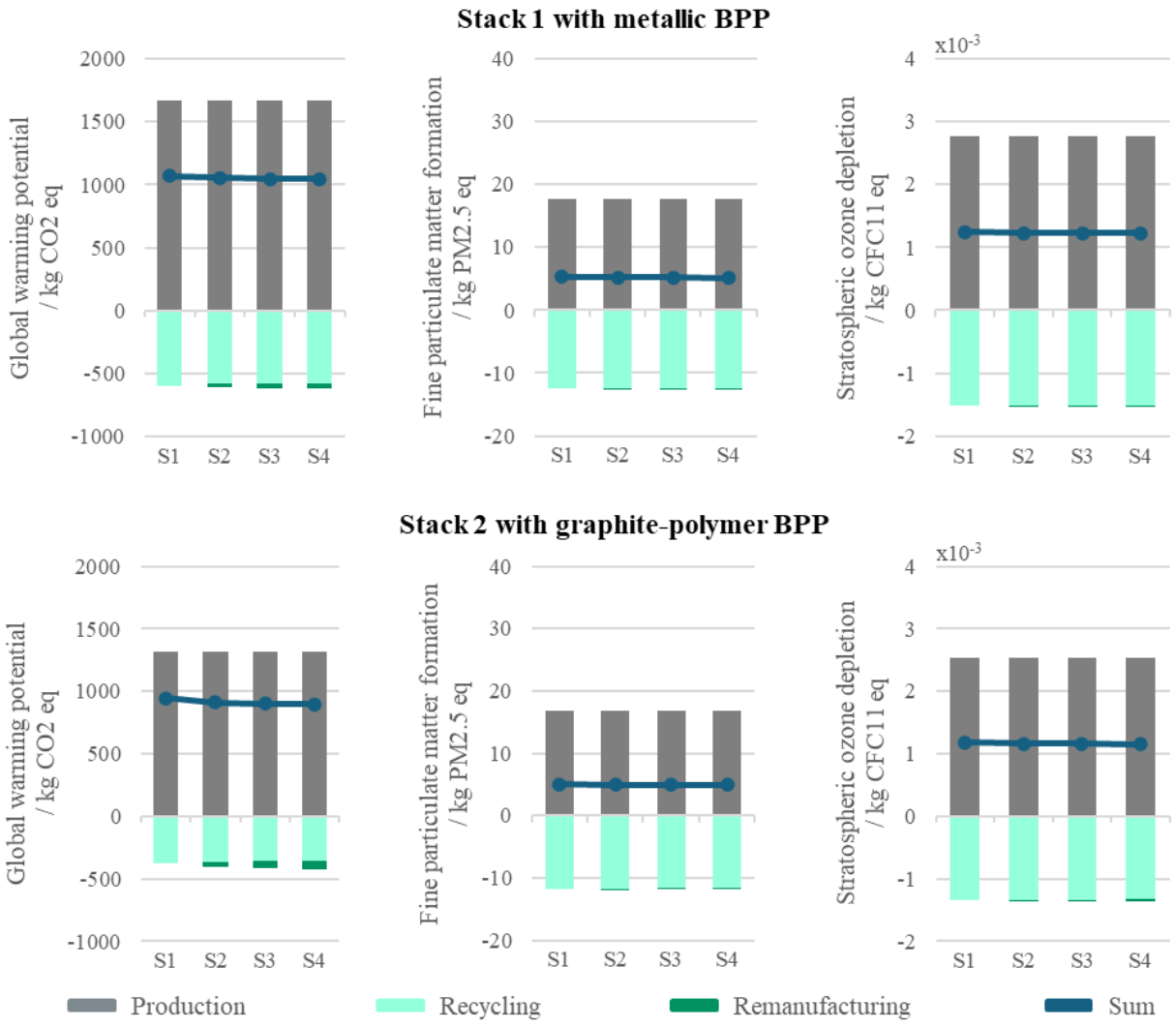


Figure 3: Scenario analysis of level I impact categories

Further analyses are limited to the impact category GWP, as this is the most common category used to assess environmental impacts.

The effect at stack level is remarkably low due to the high GWP share of the MEA, which cannot be remanufactured. But at component level, the picture changes. Table 6 lists the GWP of the different components based on the functional unit of one 55 kW PEMFC stack (both metal and graphite-polymer BPP are listed). The remanufacturable components are highlighted in grey. Column by column, the GWP of component production can be compared with the savings achieved through recycling (scenario S1) or remanufacturing (scenario S4: 90 % remanufacturing) – both in absolute and relative figures. The evaluation shows the potential that lies in the remanufacturing of components. In the case of graphite-polymer BPP, recycling can reduce the GWP of primary production by 13.0 %, while remanufacturing can reduce it by 75.7 %. For end plates, the reduction achieved through recycling is 37.6 % and through remanufacturing 92.9 %. Recycling the current collectors is already very beneficial (-88.9 %), and remanufacturing leads to a small increase to 94.4 %.

Table 6: Global warming potential of components of 55 kW stack

Component	GWP per stack				
	GWP of Production [kg CO_2 eq]	GWP of Recycling [kg CO_2 eq]	Share of Production-GWP [%]	GWP of 90 % Remanufacturing [kg CO_2 eq]	**Share of Production-GWP [%]**
MEA	1,228.8	-360.3	-29.3	-	**-**
Metal BPP	379.7	-220.8	-58.2	-	**-**
Graphite-Polymer-BPP	35.4	-4.6	-13.0	-26.8	**-75.7**
End Plate	46.3	-17.4	-37.6	-43.0	**-92.9**
Gaskets	6.7	7.6	113.4	-	**-**
Current Collectors	1.8	-1.6	-88.9	-1.7	**-94.4**
Rods, Nuts	1.3	-1.0	-76.9	-	**-**

4. Evaluation and Summary

On component level the analysis shows that the remanufacturing of BPP, end plates and current collectors reduces their ecological impact substantially. Especially for the graphite-polymer-BPP and end plates the GWP can be reduced by another 62.5 % (BPP) and 55.3 % (end plate) when compared to recycling them. Comparing the percentage saved through remanufacturing with the GWP of the total stack, this advantage becomes less noticeable: For a stack with metal BPP -2.40 % and with graphite-polymer BPP -5.08 %, cf. Table 8 and 9 in appendix. The reason for this is that the MEA causes the highest share of impacts of the stack, but from today's perspective it can only be recycled. Reusing entire cells could lead to additional impact savings. The prerequisite for this is that individual cells are diagnosed reliably and efficiently so that damage to components does not remain undetected thus compromising the performance of a remanufactured stack. Summarizing, it can be stated that the remanufacturing of PEMFC stacks and their components offers ecological advantages, and that remanufacturing strategies enabling the reuse of MEAs will have significant potential in the long term.

Acknowledgements

This work emerged from the *Hy-FCnextGen* research project and was funded by the German Federal Ministry for Economic Affairs and Energy (BMWE) based on a resolution of the German Bundestag in the funding priority "8th Energy Research Programme" with the funding code 03EN5051C.

Appendix

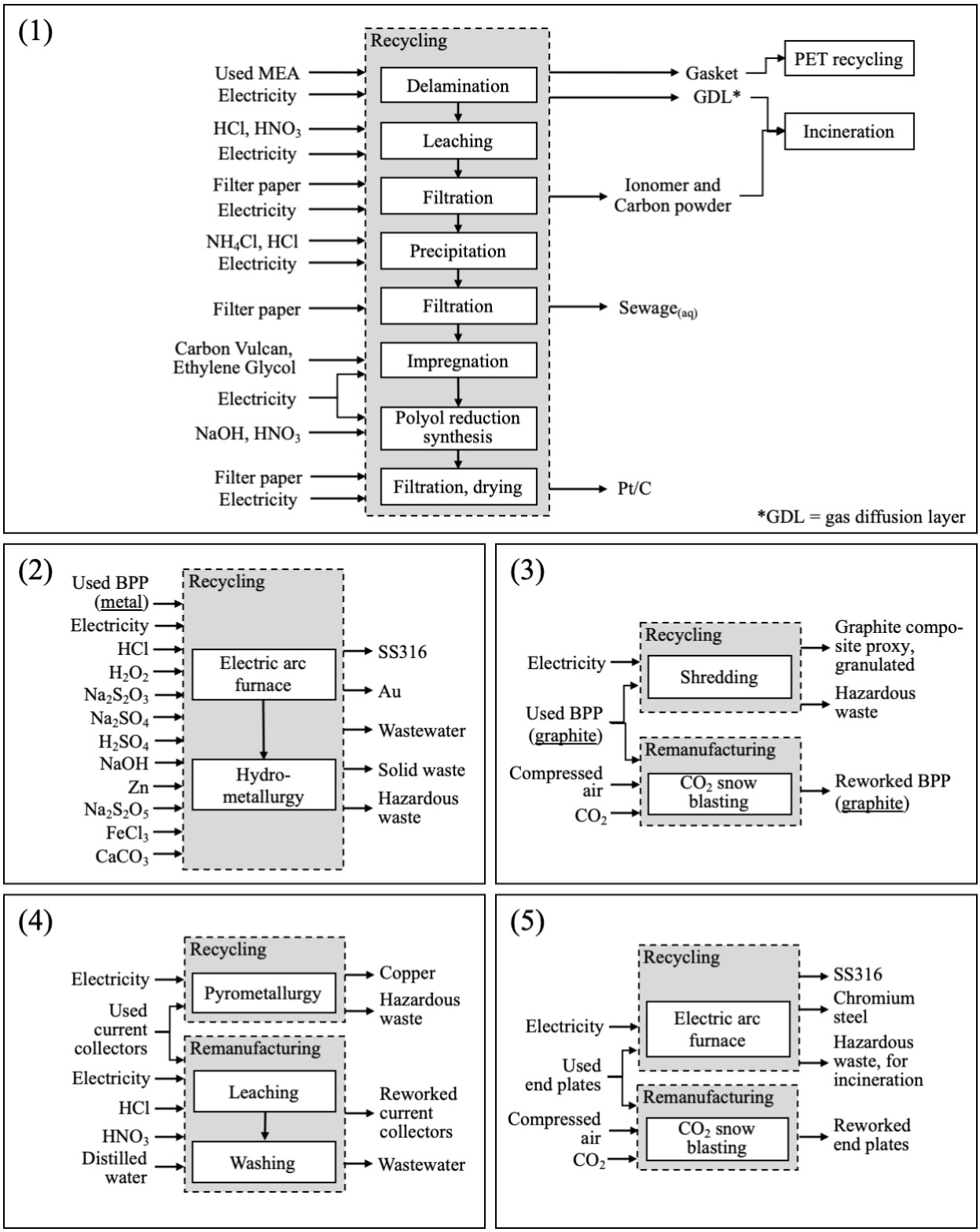


Figure 4: Flow diagram for EoL processes of MEA incl. gaskets (1), metallic BPP (2), graphite-polymer BPP (3), current collectors (4), and end plates (5)

Table 7: Life cycle inventory of the production processes of the 55 kW stack

Material or energy	Amount		Ecoinvent process used
	Stack 1 (metal)	Stack 2 (graphite)	
Inputs			
Nafion 521 Solution	393.2 g	393.2 g	Subprocesses cf. [9]
Nafion cast dispersion membrane	502.8 g	502.8 g	Subprocesses cf. [9]
PTFE	82.3 g	82.3 g	Subprocesses cf. [9]
Carbon Fiber	740.5 g	740.5 g	Subprocesses cf. [9]
Graphite	41.1 g	7,841.0 g	Graphite, battery grade {GLO}| market for | APOS, U
Pretreated Carbon	55.9 g	55.9 g	Subprocesses cf. [12]
Deionised water	35.5 g	35.5 g	Water, deionised {Europe without Switzerland}| market for water, deionised | APOS, U
Platinum	7.5 g	7.5 g	Platinum {GLO}| market for | APOS, U
Borax	11.9 g	11.9 g	Borax, anhydrous, powder {GLO}| market for | APOS, U
Methanol	35.5 g	35.5 g	Methanol {GLO}| market for | APOS, U
Carbon	47.9 g	47.9 g	Carbon black {GLO}| market for | APOS, U
Synthetic rubber	2,395.8 g	2,395.8 g	Synthetic rubber {GLO}| market for | APOS, U
Vinyl acetate	-	3,342.8 g	Vinyl acetate {GLO}| market for | APOS, U
Stainless steel (SS316)	17,921.8 g	7,171.8 g	SS316 based on Steel, chromium steel 18/8 {GLO}| market for | APOS, U
Chromium steel	366.3 g	366.3 g	Steel, chromium steel 18/8 {GLO}|market for | APOS, U
Gold	550 g	-	Gold {GLO}| market for | APOS, U
Copper	201.3 g	201.3 g	Copper, cathode {GLO}| market for | APOS, U
Glass fibre reinforced plastic	1,503.2 g	1,503.2 g	Glass fibre reinforced plastic, polyamide, injection moulded {GLO}| market for | APOS, U
Electricity	600.6 MJ	831.6 MJ	Electricity, medium voltage {Europe without Switzerland}| market group for | APOS, U
Heat	2.2 MJ	2.2 MJ	Heat, district or industrial, natural gas {RER}| market group for | APOS, U
Steam	0.07 MJ	0.07 MJ	Heat, from steam, in chemical industry {RER}| market for heat, from steam, in chemical industry | APOS, U
Sheet rolling	10,951.3 g	201.3 g	Sheet rolling, chromium steel {GLO}| market for | APOS, U Sheet rolling, copper {RoW}| processing| APOS, U
Hot rolling steel	871.2 g	871.2 g	Hot rolling, steel {Europe without Austria}| hot rolling, steel | APOS, U
Laser welding	336 m	-	Welding, arc, steel {RER}| processing | APOS, U
Waste outputs			
Hazardous waste	41.14 g	66.1 g	Hazardous waste, for incineration {Europe without Switzerland}| market for hazardous waste, for incineration |
Wastewater	34.5 l	34.5 l	Wastewater, average {Europe without Switzerland}| market for wastewater, average | APOS, U

Table 8: Impact Assessment for stack 1 (metallic BPP)

Impact category	Unit	Scenario 1	Scenario 2	Δ	Scenario 3	Δ	Scenario 4	Δ
Fine particulate matter formation	kg PM2.5 eq	5,26E+00	5,24E+00	-0,39%	5,23E+00	-0,53%	5,23E+00	**-0,59%**
Fossil resource scarcity	kg oil eq	3,16E+02	3,12E+02	-1,46%	3,10E+02	-1,94%	3,09E+02	-2,18%
Freshwater ecotoxicity	kg 1,4-DCB	1,87E+02	1,86E+02	-0,73%	1,85E+02	-0,98%	1,85E+02	-1,10%
Freshwater eutrophication	kg P eq	1,04E+00	1,03E+00	-0,68%	1,03E+00	-0,90%	1,03E+00	-1,01%
Global warming	kg CO2 eq	1,07E+03	1,05E+03	-1,60%	1,05E+03	-2,13%	1,05E+03	**-2,40%**
Human carcinogenic toxicity	kg 1,4-DCB	1,17E+02	1,14E+02	-2,50%	1,13E+02	-3,33%	1,13E+02	-3,75%
Human non-carcinogenic toxicity	kg 1,4-DCB	7,01E+03	6,99E+03	-0,28%	6,98E+03	-0,37%	6,98E+03	-0,42%
Ionizing radiation	kBq Co-60 eq	3,47E+02	3,45E+02	-0,67%	3,44E+02	-0,89%	3,44E+02	-1,00%
Land use	m2a crop eq	1,80E+02	1,79E+02	-0,19%	1,79E+02	-0,25%	1,79E+02	-0,28%
Marine ecotoxicity	kg 1,4-DCB	2,40E+02	2,39E+02	-0,73%	2,38E+02	-0,98%	2,38E+02	-1,10%
Marine eutrophication	kg N eq	7,39E-02	7,18E-02	-2,82%	7,11E-02	-3,76%	7,08E-02	-4,22%
Mineral resource scarcity	kg Cu eq	3,63E+01	3,62E+01	-0,34%	3,62E+01	-0,46%	3,61E+01	-0,52%
Ozone formation, Human health	kg NOx eq	5,34E+00	5,31E+00	-0,53%	5,30E+00	-0,71%	5,30E+00	-0,80%
Ozone formation, Terrestrial ecosystems	kg NOx eq	5,46E+00	5,43E+00	-0,54%	5,42E+00	-0,71%	5,41E+00	-0,80%
Stratospheric ozone depletion	kg CFC11 eq	1,24E-03	1,23E-03	-0,89%	1,23E-03	-1,18%	1,23E-03	**-1,33%**
Terrestrial acidification	kg SO2 eq	1,72E+01	1,71E+01	-0,29%	1,71E+01	-0,39%	1,71E+01	-0,44%
Terrestrial ecotoxicity	kg 1,4-DCB	2,49E+03	2,41E+03	-3,15%	2,38E+03	-4,20%	2,37E+03	-4,73%
Water consumption	m3	1,97E+01	1,94E+01	-1,45%	1,93E+01	-1,93%	1,93E+01	-2,17%

Table 9: Impact Assessment for stack 2 (graphite-polymer BPP)

Impact category	Unit	Scenario 1	Scenario 2	Δ	Scenario 3	Δ	Scenario 4	Δ
Fine particulate matter formation	kg PM2.5 eq	4,99E+00	4,95E+00	-0,81%	4,94E+00	-1,08%	4,93E+00	**-1,21%**
Fossil resource scarcity	kg oil eq	2,73E+02	2,65E+02	-3,20%	2,62E+02	-4,26%	2,60E+02	-4,80%
Freshwater ecotoxicity	kg 1,4-DCB	1,75E+02	1,74E+02	-0,65%	1,74E+02	-0,86%	1,73E+02	-0,97%
Freshwater eutrophication	kg P eq	9,59E-01	9,37E-01	-2,26%	9,30E-01	-3,02%	9,26E-01	-3,40%
Global warming	kg CO2 eq	9,43E+02	9,11E+02	-3,39%	9,00E+02	-4,52%	8,95E+02	**-5,08%**
Human carcinogenic toxicity	kg 1,4-DCB	8,94E+01	8,57E+01	-4,12%	8,45E+01	-5,50%	8,39E+01	-6,18%
Human non-carcinogenic toxicity	kg 1,4-DCB	6,77E+03	6,74E+03	-0,42%	6,73E+03	-0,56%	6,73E+03	-0,63%
Ionizing radiation	kBq Co-60 eq	3,15E+02	3,05E+02	-3,22%	3,02E+02	-4,29%	3,00E+02	-4,82%
Land use	m2a crop eq	1,76E+02	1,75E+02	-0,47%	1,75E+02	-0,62%	1,74E+02	-0,70%
Marine ecotoxicity	kg 1,4-DCB	2,24E+02	2,23E+02	-0,67%	2,22E+02	-0,89%	2,22E+02	-1,01%
Marine eutrophication	kg N eq	6,40E-02	6,09E-02	-4,87%	5,99E-02	-6,49%	5,94E-02	-7,30%
Mineral resource scarcity	kg Cu eq	3,55E+01	3,52E+01	-0,61%	3,52E+01	-0,81%	3,51E+01	-0,91%
Ozone formation, Human health	kg NOx eq	5,09E+00	5,03E+00	-1,08%	5,02E+00	-1,44%	5,01E+00	-1,62%
Ozone formation, Terrestrial ecosystems	kg NOx eq	5,19E+00	5,14E+00	-1,08%	5,12E+00	-1,44%	5,11E+00	-1,62%
Stratospheric ozone depletion	kg CFC11 eq	1,18E-03	1,17E-03	-1,63%	1,16E-03	-2,18%	1,16E-03	**-2,45%**
Terrestrial acidification	kg SO2 eq	1,64E+01	1,63E+01	-0,60%	1,63E+01	-0,80%	1,63E+01	-0,90%
Terrestrial ecotoxicity	kg 1,4-DCB	1,71E+03	1,69E+03	-1,07%	1,68E+03	-1,42%	1,68E+03	-1,60%
Water consumption	m3	1,38E+01	1,32E+01	-4,19%	1,30E+01	-5,59%	1,29E+01	-6,29%

References

[1] Joshi, A., Sharma, R., Baral, B., 2022. Comparative life cycle assessment of conventional combustion engine vehicle, battery electric vehicle and fuel cell electric vehicle in Nepal. Journal of Cleaner Production 379, 134407.

[2] Smit, R., Helmers, E., Schwingshackl, M., Opetnik, M., Kennedy, D., 2024. Greenhouse Gas Emissions Performance of Electric, Hydrogen and Fossil-Fuelled Freight Trucks with Uncertainty Estimates Using a Probabilistic Life-Cycle Assessment (pLCA).

[3] Mekhilef, S., Saidur, R., Safari, A., 2012. Comparative study of different fuel cell technologies. Renewable and Sustainable Energy Reviews 16 (1), 981–989.

[4] Abedin, T., Pasupuleti, J., Paw, J.K.S., Tak, Y.C., Mahmud, M., Abdullah, M.P., Nur-E-Alam, M., 2025. Proton exchange membrane fuel cells in electric vehicles: Innovations, challenges, and pathways to sustainability. Journal of Power Sources 640, 236769.

[5] Hart, D., Jones, S., Houghton, T., Cordobes, X., Lewis, J., 2021. The Fuel Cell Industry Review 2021.

[6] Weichenhain, U., Lange, S., Koolen, J., Benz, A., Hartmann, S., Heilert, D., Henninger, S., Kallenbach, T., 2020. Potenziale der Wasserstoff- und Brennstoffzellen-Industrie in Baden-Württemberg.

[7] Al Assadi, A., Goes, D., Baazouzi, S., Staudacher, M., Malczyk, P., Kraus, W., Nägele, F., Huber, M.F., Fleischer, J., Peuker, U., Birke, K.P., 2023. Challenges and prospects of automated disassembly of fuel cells for a circular economy. Resources, Conservation & Recycling Advances 19, 1–15.

[8] Schmidt, P.A., 2022. The Contribution of new Production Technologies and Circular Economy Towards meeting the Future Demand of Proton-exchange Membrane Fuel Cells – A Literature Review. Hannover : publish-Ing, 10 pp.

[9] Evangelisti, S., Tagliaferri, C., Brett, D.J., Lettieri, P., 2017. Life cycle assessment of a polymer electrolyte membrane fuel cell system for passenger vehicles. Journal of Cleaner Production 142 (4), 4339–4355.

[10] Iannuzzi, L., Hilbert, J.A., Silva Lora, E.E., 2021. Life Cycle Assessment (LCA) for use on renewable sourced hydrogen fuel cell buses vs diesel engines buses in the city of Rosario, Argentina. International Journal of Hydrogen Energy 46 (57), 29694–29705.

[11] Simons, A., Bauer, C., 2015. A life-cycle perspective on automotive fuel cells. Applied Energy 157, 884–896.

[12] Gottardo, G., Peressut, A.B., Colnago, S., Latorrata, S., Piegari, L., Dotelli, G., 2023. LCA of a Proton Exchange Membrane Fuel Cell Electric Vehicle Considering Different Power System Architectures. Energies 16 (19), 6782.

[13] Gulotta, T.M., Salomone, R., Mondello, G., Saija, G., Lanuzza, F., Briguglio, N., 2023. Life Cycle Assessment and Environmental Life Cycle costing of a unitised regenerative fuel cell stack. The Science of the total environment 901, 166007.

[14] Mori, M., Iribarren, D., Cren, J., Cor, E., Lotrič, A., Gramc, J., Drobnič, B., Rey, L., Campos-Carriedo, F., Puig-Samper, G., Bargiacchi, E., Dufour, J., Stropnik, R., 2023. Life cycle sustainability assessment of a proton exchange membrane fuel cell technology for ecodesign purposes. International Journal of Hydrogen Energy 48 (99), 39673–39689.

[15] Pehnt, M., 2001. Life-cycle assessment of fuel cell stacks. International Journal of Hydrogen Energy 26 (1), 91–101.

[16] Granados-Fernández, R., Montiel, M.A., Díaz-Abad, S., Rodrigo, M.A., Lobato, J., 2021. Platinum Recovery Techniques for a Circular Economy. Catalysts 11 (8), 937.

[17] Duclos, L., Svecova, L., Laforest, V., Mandil, G., Thivel, P.-X., 2016. Process development and optimization for platinum recovery from PEM fuel cell catalyst. Hydrometallurgy 160, 79–89.

[18] Duclos, L., Chattot, R., Dubau, L., Thivel, P.-X., Mandil, G., Laforest, V., Bolloli, M., Vincent, R., Svecova, L., 2020. Closing the loop: life cycle assessment and optimization of a PEMFC platinum-based catalyst recycling process. Green Chemistry 22 (6), 1919–1933.

[19] Stropnik, R., Mori, M., Lotrič, A., Gramc, J., 2023. D5.1. Environmental profile of existing EoL technologies and effects in the scope of circular economy in manufacturing phase Ref. Ares(2023)4533141 - 30/06/2023. Best4Hy. https://best4hy-project.eu/wp-content/uploads/2023/09/BEST4Hy_D5.1-Environmental-profile-of-EoL-tech.pdf. Accessed 21 February 2024.

[20] Oldewurtel, S., Wagner, A., Imdahl, C., Zellmer, S., Dröder, K., 2024. Disassembly technologies for PEMFC stacks in heavy-duty applications. Procedia CIRP 127, 218–223.

[21] Goes, D., 2024. Projektbeschreibung ReStack: Remanufacturing von PEM-Brennstoffzellenstacks für eine nachhaltige Kreislaufwirtschaft.

https://www.wbk.kit.edu/wbkintern/Forschung/Projekte/ReStack/index.php?site=home. Accessed 23 March 2025.

[22] Deutsches Institut für Normung e.V., 2009. DIN EN ISO 14040:2006-11. Umweltmanagement - Ökobilanz - Grundsätze und Rahmenbedingungen. Beuth Verlag, Berlin.

[23] Deutsches Institut für Normung e.V., 2021. DIN EN ISO 14044:2021-02. Umweltmanagement - Ökobilanz - Anforderungen und Anleitungen. Beuth Verlag, Berlin.

[24] Stropnik, R., Mori, M., Lotrič, A., Gramc, J., 2023. BEST4Hy-D5.1_Environmental profile of existing EoL technologies and effects in the scope of cirular economy in manufacturing phase.

[25] Gautam, R., Banerjee, S., Kar, K., 2015. Bipolar Plate Materials for Proton Exchange Membrane Fuel Cell Application. Recent Patents on Materials Science 8 (1), 15–45.

[26] EKPO Fuel Cell Technologies GmbH, 2023. PEMFC stack module NM5-EVO. https://www.ekpo-fuelcell.com/en/products-technology/fuel-cell-stacks. Accessed 18 March 2024.

[27] Duclos, L., Lupsea, M., Mandil, G., Svecova, L., Thivel, P.-X., Laforest, V., 2017. Environmental assessment of proton exchange membrane fuel cell platinum catalyst recycling. Journal of Cleaner Production 142, 2618–2628.

[28] Mori, M., Lotrič, A., Stropnik, R., 2019. D4.3 Case studies with new strategies in dismantling and recycling stage. Grant Agreement No. 700190. HyTechCycling. https://hytechcycling.eu/wp-content/uploads/d4-3-case-studies-with-new-strategies-in-dismantling-and-recycling-stage.pdf. Accessed 29 March 2024.

[29] Zhu, P., Driscoll, E.H., Dong, B., Sommerville, R., Zorin, A., Slater, P.R., Kendrick, E., 2023. Direct reuse of aluminium and copper current collectors from spent lithium-ion batteries. Green Chemistry 25 (9), 3503–3514.

[30] Kranert, M., 2017. Einführung in die Kreislaufwirtschaft. Planung - Recht - Verfahren, 5. Auflage ed. Springer Vieweg, Wiesbaden.

[31] Ahlfs, S., Goudz, A., Streichfuss, M., 2020. Die Brennstoffzelle. Springer Fachmedien Wiesbaden, Wiesbaden, 65 pp.

[32] Li, X., 2023. Bipolar plates and flow field design, in: Das, P.K., Jiao, K., Wang, Y., Barbir, F., Li, X. (Eds.), Fuel Cells for Transportation. Fundamental Principles and Applications. Elsevier, Amsterdam, pp. 305–337.

[33] IAG Industrie Automatisierungsgesellschaft m.b.H., 2025. Nachbehandlung graphitische Bipolarplatten | IAG. https://www.iag.at/nachbehandlung-graphitische-bipolarplatten#. Accessed 29 May 2025.

[34] Schulz, D., 2021. Wie lassen sich Bipolarplatten am besten reinigen? JOT Journal für Oberflächentechnik 61 (9), 62–65.

[35] Stahl, H., Bauknecht, D., Hermann, A., Jenseit, W., Köhler, A.R., Merz, C., Möller, M., Schüler, D., Jörissen, L., Storr, U., 2016. Ableitung von Recycling- und Umweltanforderungen und Strategien zur Vermeidung von Versorgungsrisiken bei innovativen Energiespeichern. Umweltbundesamt, Dessau-Roßlau.

[36] Schmidt, M., Haubach, C., Schäfer, P.R., Schebek, L., Zeller, V., Goellner, L., Hagedorn, T., Haas, S., Neitz-Regett, A., Nill, M., Drinkuth, H., Ding, Y., 2022. Entwicklung einer standardisierten Vorgehensweise zur Ermittlung der eingesparten Treibhausgasemissionen aus Maßnahmen zur Materialeffizienz (ESTEM). Abschlussbericht. VDI Zentrum Ressourceneffizienz, Berlin. Accessed 11 August 2024.

[37] European Commission-Joint Research Centre - Institute for Environment and Sustainability, 2011. International Reference Life Cycle Data System (ILCD) Handbook: Recommendations for Life Cycle Impact Assessment in the European context, First edition ed. Publications Office of the European Union, Luxembourg, 1143 pp.

Biography

Marie Beller (*1993) has been a Scientific Researcher at the Chair of Manufacturing and Remanufacturing of the University of Bayreuth since 2021. Her research focus is investigating the feasibility of circular economy strategies – for fuel cells in particular. She holds a M.Sc. in Management and Technology from the Technical University of Munich (TUM).

Manuel Querbach (*1999) studied industrial engineering at the University of Bayreuth and completed his master's degree in 2024 with a thesis at the Chair of Manufacturing and Remanufacturing addressing the ecological benefits of remanufacturing processes. This thesis, which focused on circular economy strategies for fuel cells, provides the foundation for the research presented in this paper.

Frank Döpper (*1970) holds the Chair of Manufacturing and Remanufacturing at the University of Bayreuth and is scientific director "Manufacturing and Process Engineering" at Fraunhofer IPA. After studying mechanical engineering at RWTH Aachen University, he began his career at the Fraunhofer IPT in Aachen before moving into industry. For more than twenty years, he worked in technology and investment management and as a site and production manager for international companies before he was appointed professor at the University of Bayreuth.

8th Conference on Production Systems and Logistics

Towards Explainable Knowledge And Behavior In Social Robotics Using Fuzzy Techniques: A Systematic Literature Review

Raoul Zebisch, Johannes Schilp
Chair of Digital Manufacturing, Faculty of Applied Computer Science, University of Augsburg, Augsburg, Germany

Abstract

In recent years, research in social robotics has gained popularity in several application fields, most prominently healthcare. Enhanced human-robot interaction provides added value in elderly care, where social capabilities can improve the quality of life. Likewise, the topic has also become relevant in other fields, such as manufacturing. A deployed robot's social skills can increase the perceived comfort of collaborative tasks, such as resolving machine malfunctions or training new workers, improving the overall productivity and robot acceptance in modern factories. However, their validation through digital approaches proves to be unarguably difficult. Without a formal tool to encode, store, and retrieve social knowledge about scenarios of industrial human-robot interaction, there is no way for a computer to tell how a complex behavior model with a multitude of social abilities performs in terms of human trust, user acceptance, or naturalness of an interaction. Due to the imprecise nature of social facts and knowledge, encoding this information is a nontrivial task. Mathematical models based on fuzzy logic and fuzzy sets have the potential to express a social robot's behavior and cognition. Therefore, the authors propose a prospective approach with corresponding design goals towards an explainable social knowledge base through semantic formalization and fuzzy techniques as an initial step to provide said tool for social robot validation as an explainable rule-based system. An extensive and reproducible systematic literature review on applications of fuzzy techniques in social robotics supports the argumentation. The used review scheme is PRISMA2020.

Keywords

Social Robots; Fuzzy Logic; Human-Robot Interaction; Approximate Reasoning; Fuzzy Semantics

1. Introduction

In recent years, social robotics has gained high scientific interest. Modern breakthroughs in artificial intelligence promise new possibilities in this field. A social robot is a system that behaves in a way during interaction that feels natural and human-like [1]. So far, researchers and companies deploy social robots mostly outside manufacturing, in fields such as service, entertainment, education, healthcare, and research [2]. Closing this gap makes foundational research towards virtual validation of explainable social robotic behavior more relevant. From a computer scientist's point of view, creating algorithmic solutions to rate a robot's social capability is a daunting task. Under certain circumstances, a multitude of social facts [3] need to be considered during human-robot interaction and formalized into an information model that a computer can understand to provide an answer to the question of how well the system likely performs in a specific use case.

In digital manufacturing, validation of production processes usually happens through simulation. A recent trend is the deployment of digital twins to represent assets (like machines, robots, or persons) and entire process chains through a digital replica [4]. In some cases, these applications feature semantic knowledge

DOI: https://doi.org/10.15488/20970
ISSN: 2701-6277

bases that provide domain-specific information on the simulated entities. These systems have an additional layer of formalization applicable to processes revolving around social human-robot interaction. However, ontological approaches based on RDF or OWL [5] are limited in their expressiveness, since they can only store crisp logical propositions and rules.

The authors, therefore, became interested in a method that incorporates imprecision to semantic domain knowledge, thus focusing on fuzzy logic [6] as a theoretical foundation. By imprecision, the authors mean that linguistic terms describing social behavior, such as a robot's perceived friendliness (very friendly, friendly, neutral, rude, very rude), do not have a crisp logical definition and can only be expressed through imprecise statements and quantifiers, i.e., fuzzy logic. Fuzzy knowledge graphs as data models to encode approximate social facts seem like a promising fit for the formalization of social robots and their interaction processes with humans. Through social queries to this fuzzy semantic knowledge base, a robot could retrieve and infer social facts that aid in the decision process for adequate social behavior. Such a knowledge base needs consideration, which applications in social robotics are suitable for fuzzy techniques, and what kind of social capability might be expressible through them, leading to the following three research questions:

- **RQ 1**: How are fuzzy techniques already utilized in social robotics?
- **RQ 2**: What kind of social robot capability is modelled through fuzzy techniques?
- **RQ 3**: How can a fuzzy-semantic knowledge base for social robots be implemented?

This paper has two main contributions:

- A systematic and reproducible literature review based on PRISMA2020 [7] on the application of fuzzy techniques in social robotics regarding works relevant to one or more social abilities to answer RQ 1 and RQ 2.
- A prospective proposal of a fuzzy-semantic database and query engine for imprecise social knowledge bases to answer RQ 3. For clarification, the authors also provide an example use case backed by the findings of the conducted literature review.

The structure of this paper is as follows: The remainder of Section 1 will be a theoretical summary of fuzzy logic and formal semantics. Section 2 discusses the findings of the systematic literature review. Section 3 explains future work on a prospective approach towards a formalization tool of social robots using fuzzy semantic knowledge graphs clarified through an example scenario regarding robot proxemic zones. Finally, Section 4 concludes the paper.

1.1 Fuzzy logic and fuzzy set theory

Fuzzy logic, as introduced by Zadeh [6], is a superset of Boolean logic. Unlike regular logic, which only differentiates in statements that are either *true* or *false*, fuzzy logic adds to each element in a set of logical propositions P a truth value $T(p)$ with $p \in P;\ T(p) \in [0,1]$. Precise logic would be a specialization with $\forall p, T(p) \in \{0,1\}$. Fuzzy logic contains generalized versions of the traditional Boolean operators, such as:

$$T(p_1 \wedge p_2) := \min[T(p_1), T(p_2)] \tag{1}$$

$$T(p_1 \vee p_2) := \max[T(p_1), T(p_2)] \tag{2}$$

$$T(p_1 \rightarrow p_2) := \max[\min[T(p_1), T(p_2)], 1 - T(p_1)] \tag{3}$$

Fuzzy logic offers the possibility to define custom quantifiers, such as *most, many, several, few, occasionally*, which opens the possibility to describe linguistic variables with semantic meaning. An example would be the linguistic variable of student lecture presence, which takes on the value *high* if the student attends *most* lectures. In the same sense, fuzzy set theory is a superset of regular set theory in which elements x of a fuzzy set X can have partial membership $\mu(x)$. The membership of elements for set union and intersection behaves analogous to their logical counterparts of conjunction and disjunction.

1.2 Semantic knowledge graphs

Semantic knowledge graphs are a form of data model building on the principles of first order logic. A common variant would be the Resource Description Framework [5], which encodes knowledge as a set of triples that have the form (s, p, o), with a subject s, a predicate p and an object o. An example of a triple expressing the fact *Augsburg is a city* could be encoded in RDF as follows:

$$< https://www.augsburg.de\#uri > \ rdf:type < https://www.example.com/City > .$$

Subject and predicate are globally identifiable vocabulary in form of unified resource identifiers (URI), while the object o is either a URI or a literal value. The connection between regular logic and formal semantics is the fact that each triple of an RDF graph is essentially a logical proposition. Therefore, as with fuzzy logic, one can add truth values to each triple resulting in fuzzy semantic knowledge graphs.

2. Review on fuzzy techniques in social robotics

The authors conducted a systematic literature review to research the state-of-the-art of fuzzy technique applications in social robotics. The methods described in each scientific work use fuzzy reasoning to achieve some form of social robot capability. The authors define social capability as a set of social abilities a robot or human can express. Unfortunately, the terms "social ability" and "social skill" are not well-defined, which is why this paper provides a set of definitions of the ones identified in the literature review. On the one hand, a social ability in human-robot interaction scenarios can be present even though the work covering it does not have a clear focus on social robotics. On the other hand, one can argue that as soon as there is an interaction scenario between humans and robots, a social aspect is always attached to it. Therefore, the authors hope that the scientific consensus can agree on the provided definitions of a robot's intuitive social abilities.

2.1 Review process

The literature review follows the PRISMA2020 scheme [7]. Its results are reproducible using the files stored in the repository at *https://github.com/Neroware/CPSL2026-Reproduce-SLR*. The search strings remained the same for all databases/registers, using only simple keywords and disjunction/conjunction operators:

- **Q1**: ("social robotics" OR "social robot" OR "companion robot") AND ("fuzzy logic" OR "fuzzy inference" OR "fuzzy control" OR "fuzzy system")
- **Q2**: ("social robotics" OR "human-robot interaction" OR "socially assistive robot") AND ("fuzzy semantics" OR "fuzzy linguistics" OR "linguistic variables" OR "fuzzy reasoning" OR "approximate reasoning")
- **Q3**: ("human-robot interaction") AND ("emotion" OR "behavior modeling" OR "affective computing") AND ("fuzzy logic" OR "fuzzy inference" OR "fuzzy sets" OR "fuzzy control" OR "fuzzy system")

The authors retrieved all records up to 02nd July 2025. The Springer Link database is an exception. Unfortunately, it is impossible to extract both the abstract and title from its website. Therefore, search queries went to the Springer Metadata API using a scripted solution. In this case, the publication date lies between 2019 and 2024. A brief argumentation for the individual search strings follows:

- **Q1**: A search string specifically for fuzzy logic and control in social robotics
- **Q2**: A search string for broader fuzzy reasoning with a focus on semantic and linguistic variables
- **Q3**: A search string for papers with no direct focus on social robotics but with relevancy due to aspects like behavior or emotion modelling in human-robot interaction using fuzzy inference

The PRISMA2020 flow chart can be found in Appendix 1. After the screening process a total of 54 studies were included in the review.

2.2 Identified social abilities

This subsection provides definitions of the social abilities the robots express in the reviewed literature.

Social-cognitive movement and positioning: The ability to move socially acceptable. A robot that moves under social-cognitive considerations takes human safety and movement intention into account but also maintains socially acceptable positioning and distances. One modelling option for the latter would be human proxemic zones as described in [8], [9], or [10]. Human proxemic zones distinguish between multiple areas, such as a public zone, a social zone for interaction, a personal zone, and an intimate zone. Fuzzy inference can determine whether a robot maintains an acceptable social distance to humans by stating fuzzy rules for proxemic boundaries, as in [8].

Human emotion and intention recognition: The ability to detect a person's emotional state and/or human intention. While both aspects differ, the authors decided to group them under one social ability class, as in most reviewed literature, such as [11], [12], and [13], the detection of the emotional state was essential in determining an intention assumption. The various studied works associated with Prof. Hirota [11, 14-21] rely on emotion recognition using machine learning approaches combined with fuzzy rules and inference to determine an assumed intention. Other aspects, such as predicting human movement intent [13], also fall under this category, featuring fuzzy inference engines.

Simulation and expression of artificial emotions: The ability to simulate and/or express an emotional state. An empathic robot would have the ability not only to detect but also to enact adaptive responses to perceived human emotions. Emotional expressiveness can occur, among others, through a humanoid robot's face, as in [22], or through expressive gestures as shown with the robotic arm in [23]. Fuzzy logic provides rules for transitions between emotional states, as in [24] or [25]. Additionally, this category is relevant for programming robot behavior via rule-based approaches [25].

Social-cognitive communication based on domain-specific knowledge: The ability to react to or engage in human communication utilizing domain-specific knowledge, such as social context or communication atmosphere. An example is the linguistic understanding of user commands, as in [26]. Depending on the social context, lexical terms, such as the word pen, can have different semantic meanings. In this case, the word can either be a writing tool or a female swan. The robot utilizes domain knowledge and fuzzy sets to associate lexical terms with a valid social context. In [9], the robot adapts its proxemic zones accordingly to communicated user feedback using linguistic variables and fuzzy inference. In [27], the robot uses domain knowledge and fuzzy inference to provide adequate compliments to a human user.

Artistic co-creation: A robot's ability to engage in creative processes with humans. In the literature review, the authors identified the work by Thörn, Knudsen, and Saffiotti, which covers a collaborative robotic system that dances accordingly to an improvising human music performer. It was the only work that fell under this specific category of social capability. [28]

Person recognition: The ability to recognize and identify human persons. The work of Siow et al. uses computer vision networks and a fuzzy model to fuse observations of human faces and poses into a final recognition result [10]. Al-Qaderi, Lahamer, and Rad make use of fuzzy rules to fuse the outputs of multiple classifiers used for human speaker recognition into a single prediction [29].

Social-cognitive admittance control: The ability to accordingly adjust mechanical forces in scenarios where humans and robots are in physical contact. Admittance control is relevant, especially in research on rehabilitation robots. The two works [30, 31] from the literature review do not directly take social considerations into account, but make use of fuzzy inference mechanisms nonetheless, which could be expanded by social-cognitive considerations. Since the authors argued that a social aspect is always present in any human-robot interaction, they decided to feature these two studies for completeness.

Collaborative assistance and division of labor: The ability to recognize in a scenario of human-robot interaction, which tasks are assignable to the robot, and which to the human. The robot can take human safety and ergonomics considerations into account, as in [32]. While not directly aimed at social robotics, the reviewers argue that this is a form of social ability, nonetheless, due to its strong natural association with the terms "helpfulness" or "cooperativeness".

2.3 Discussion

This section discusses the review's most relevant findings. Appendix 2 summarizes the results of the literature review per work and contains the following information:

- A reviewed work's reference
- A listing of the social abilities the work is related to
- A summary of the fuzzy technique used by the work's authors
- A possible social fact query, which corresponds to the robot's social-cognitive considerations

Fuzzy machine learning for emotional intention estimation: As stated in Section 2.2, the authors decided to group intention and emotion recognition under one social ability. While other aspects, such as movement or interaction intent, also fall under this category, 11 reviewed papers specifically refer to the recognition of emotional intention through fuzzy machine learning techniques. In these cases, the robot can learn human intention based on human emotional cues, complemented with additional personal information. Most prominently are the works associated with Prof. Hirota, which account for 9 contributed records in the review. Their work heavily focuses on emotional intention recognition using computer vision and personal user information to determine possible preferences, such as drinking preferences in Japanese-style Bars (ja: 居酒屋). One of their goals is to make service robots deployed as bartenders more customer-friendly by enabling them to give suggestions to guests when ordering [11]. Initially working with fuzzy Q-learning to infer the communication atmosphere in social environments [15], their work includes architectures based on Fuzzy Support Vector Regression with training data clustered through Fuzzy C-Means [11, 16]. Their Two-Layered Fuzzy Support Vector Regression (TLFSVR) model is deployed in an intention understanding architecture, which makes use of users' personal data and emotion recognition from facial expression to acquire a final intention prediction, and consists of four modules: An ID mapping module to distinguish between familiar and unfamiliar users, a fuzzy rule base to retrieve intentions directly from memory in case of a familiar person, the TLFSVR to learn the relationship between recognized emotions and the user's identification information, and lastly a fuzzy fusion model using fuzzy rules to determine the predicted intention. [11] In more recent works, they also utilized fuzzy deep learning techniques such as a Fuzzy Deep Neural Network with Sparse Autoencoder [20] or a broad deep emotional intention inference model based on convolutional neural networks and fuzzy inference models [19]. In a similar setup, Hao, Cao, and Liu propose a fuzzy model architecture to determine service demands for a robot supplying water to users. Their method utilizes a Takagi-Sugeno model to infer the degree of human intention as a linguistic variable, enabling the robot to provide its service. [12] Finally, the architecture proposed by Bien et al. aims to enable robots to autonomously provide services in smart houses as an option for healthcare at home. Their system utilizes fuzzy rules to determine human intention from hand gestures modelled as a linguistic variable. According to the authors, the same method would apply to the expression of human facial emotions. [33]

Fuzzy rules for expression of social behavior, such as emotional states: Behavior programming with fuzzy rules is explainable and predictable, making it viable in scenarios where robots need to express complex social abilities, such as an internal emotional state. A social robot's ability to simulate and express human emotions is fundamentally different than just recognizing it from peers. Furthermore, one might argue that behavior programming through fuzzy logic is unrelated to human emotions. However, in the reviewed literature, behavior programming is either used when modelling emotional expressive states, or to express a robot's cognition of social behavior, which also prominently features emotional states. The works

contributed by [24] and [25] model *Fear* or *Anxiety*. Interestingly, these two works both cover examples of simulated animal behavior, such as a dog missing its owner, which is a simulated emotional response. Both works use state machines with Mamdani fuzzy rules for transition. Likewise, [34] and [35] make use of this method to simulate emotional expressive states, i.e., for a robot's body emotion regarding the first work, and a robot's facial expression regarding the second. The robot platform introduced by Ponce, Molina, and Grammatikou [36, 37] and the kinematic arm with expressive gestures introduced by [23] both feature expressive states, though, focus is more on determining an expressive state directly from observed sensor input using fuzzy inference than on utilizing a concrete fuzzy behavior model.

Linguistic understanding through fuzzy semantics: When humans try to communicate with robots, e.g., through voice utterance, there needs to be mutual understanding between both sides. Social cognitive communication is essential for a robot to feel natural, even though a clear linguistic and semantic understanding is challenging for a machine. Communication and resulting responses vary drastically depending on the context, which is why additional domain-specific knowledge is required. This knowledge consists of information, social facts, and rules specific to the robot's social context. A prominent example is the work of Jusoh and Alfawareh on linguistic understanding. They propose a semantic database for context-sensitive communication, which provides the semantic meanings of vocabulary terms to a robot using fuzzy set theory. [26] This way, ambiguous expressions can be linked to their specific context, which is essential in high-context-sensitive languages such as Japanese. (There are over 20 different meanings associated with the word "kouka" (ja: こうか)). Vincze and Niitsuma use fuzzy rules and natural language processing to enable a robot for parent-child interaction therapy to extract information on the communication atmosphere (linguistic variables of happiness or sadness) [38]. Furthermore, [39] and [40] focus on the linguistic meaning of movement commands, which also depend on the robot's context, such as physical location. Using Mamdani inference, semantic meaning and connected movement actions are put into relation.

Fuzzy robot controllers: Fuzzy techniques have high applicability in control engineering. It is therefore unsurprising that they are suitable for human-aware robot controllers, essential for socially aware movement or admittance control. The authors of [41] propose a fuzzy robot controller, modelling ultrasonic sensor input as linguistic variables to provide human-aware obstacle avoidance at close proximities. The authors of [42] propose a social force model with a Mamdani inference to determine the linguistic variable of the robot's linear velocity. The authors of [43] chose a similar approach, utilizing Mamdani inference to deduce linear and angular velocity as linguistic variables. Fuzzy controllers are also deployed by [30] for admittance control of a rehabilitation robot to determine the end-effector velocity as a linguistic variable. In their specific case, they complement their fuzzy rules with additional semantic depth through hedge algebra. The authors of [31] deploy a Takagi-Sugeno model for their fuzzy controller to determine the robot's virtual damping value.

Additional contributions: Thörn, Knudsen, and Saffiotti propose a robotic system of artistic co-creation that can improvise dance moves based on the music played by a human performer. The system uses fuzzy inference to determine the complexity of the robot's movements. [28] Al-Qaderi, Lahamer, and Rad propose a heterogeneous classifier architecture for person recognition based on a user's utterance, utilizing fuzzy inference for model output fusion [29]. Likewise, Siow et al. propose a fuzzy classifier model based on Takagi-Sugeno inference fusing body and face detections for person recognition [10]. Alession, Aliev, and Antonelli propose a fuzzy inference system for collaborative assembly using Mamdani fuzzy rules to determine whether the robot or the human worker can perform a task. One example linguistic variable in their work is a fuzzified NIOSH index, which is relevant in ergonomics. [32] Finally, Dell'Anna and Jamshidnejad propose SONAR, a robot aware of social norms, utilizing a fuzzy inference model. Their contribution is particularly relevant, since membership functions are continuously optimized. [44]

3. Future work on formalized explainable social knowledge in robotics

In this section, the authors present a prospective semantic database system and query processor. Its objective is to query and retrieve imprecise social knowledge by leveraging fuzzy techniques and semantic web technologies such as OWL/RDF. Based on the reviewed works, the authors identified the following design goals:

- **Goal 1**: The database relates the semantic meaning to the stored data
- **Goal 2**: The database can store imprecise logical propositions
- **Goal 3**: The data processor performs logical inference on query processing
- **Goal 4**: If possible, query processing and inference happen in real-time

The purpose of this knowledge base is to enable robots to query social facts, which is why the authors emphasize the social questions identified in the reviewed literature that robots must consider (social queries). Goal 1 is relevant to connecting linguistic variables in imprecise inference engines with their semantic meanings, which are essential for linguistic understanding. Goal 2 ensures that imprecise social knowledge is formalized and stored in a social knowledge base. Goal 3 allows deductive processes, enabling the system to answer a robot's social queries. Finally, Goal 4 ensures that these deductive processes can occur during social interaction between humans and robots.

3.1 Example use case: Proxemic zones as social fact query

For reader comprehension and as an outlook towards future research, the authors provide a basic example use case that utilizes the prospective social knowledge base and data processor. This scenario is loosely related to [8] and describes a robot's social ability of social-cognitive movement and positioning. In this use case, the robot adjusts proxemic zones around a person according to perceived social parameters, such as facially expressed emotions, estimated age, and the distance between the robot and the individual. It is difficult to deny that a robot's ability to increase distance and reduce movement speed around individuals who are children or appear frightened constitutes a social capability. In such a scenario, the input and output parameters are describable as linguistic variables of a Mamdani inference system with fuzzy rules such as:

- IF Age = *Very Young* AND Facial Expression = *Scared* AND Distance = *Medium* THEN Proxemic Zone = *Personal* AND Movement Speed = *Very Low*.
- IF Age = *Young* AND Facial Expression = *Happy* AND Distance = *Medium* THEN Proxemic Zone = *Social* AND Movement Speed = *Low*.

That is to say, the proxemic zones shift depending on the observed emotion, and the movement speed is adjusted if the person is in a vulnerable age group. Then, one can phrase the robot's social query as: *What is Robot X's proxemic positioning and maximum movement speed, if a Person Y's observed facial expression has emotion vector V_E, Person Y's observed age is n years, and the distance between Robot X and Person Y is d meters?* Figure 1 visualizes the scenario.

3.2 Thoughts on implementation

Regarding the previously stated design goals and the social query from Section 3.1, one possibility to implement the described system is the extension of RDF and its associated query language SPARQL by fuzzy semantic features such as linguistic variables, fuzzy inference rules, and truth values, which automatically serve Goal 1 and Goal 2 (Section 1.1 & 1.2). In doing so, linguistic variables and fuzzy rules receive semantic meaning through vocabulary provided in ontological models. Furthermore, the semantic meaning of entities like the emotion vector V_E is maintained and reusable in other social abilities, such as emotional intention prediction. To achieve Goal 3, which enables the deduction of new fuzzy knowledge, the semantic database would also require membership functions to be formalized and stored within it. An example is the fuzzification process of age as a linguistic variable. In the given case, the observation on a person's age can be stored and fuzzified as the following quadruples:

$$(ex{:}\,PersonA; rdf{:}\,type; ex{:}\,Person; 1.0); (ex{:}\,PersonA; ex{:}\,age; 10\text{^^}xsd{:}\,long; 1.0);$$

$$(ex{:}\,PersonA; ex{:}\,age\,; ex{:}\,Young; 0.67);\ (ex{:}\,PersonA; ex{:}\,age\,; ex{:}\,VeryYoung; 0.33)\ .$$

Note that the fourth tuple component defines the membership of the RDF-triple. If the semantic database also encodes and stores the membership function, the third and fourth quadruples are deductible. In this example, fuzzification resulted in the arbitrarily selected values of 0.67 and 0.33, respectively.

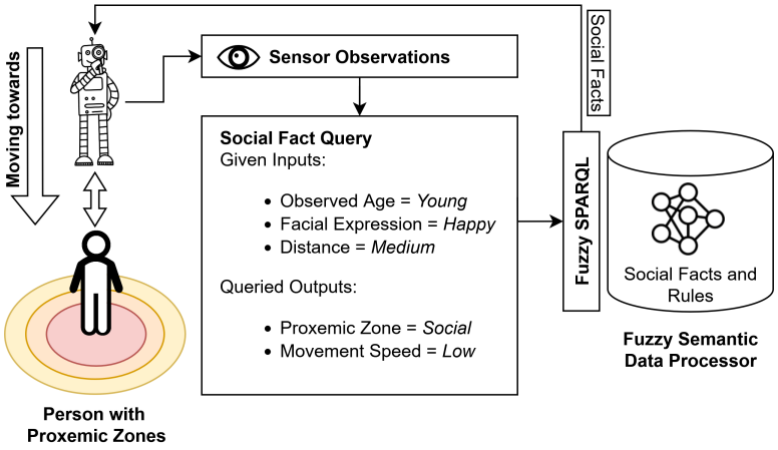


Figure 1: Use case of a robot's social-cognitive movement with fuzzy social knowledge base

Quadruples and membership functions would be stored in a Property Graph, as used by the graph database system Neo4J, for fast node-edge traversal serving Goal 4. Achieving efficient processing of social queries formalized in a fuzzy SPARQL superset would be possible through a just-in-time compiled data processor with optimizations on query sub-operator level [45]. Figure 2 visualizes the prospective query processing pipeline. The system would formulate social fact queries with linguistic variables in the SPARQL superset frontend, translating them into an algebra tree, which is then lowered to sub-operators enabling direct optimizations on graph data structures. Lastly, the query engine translates the declarative sub-operator plan to an imperative platform-specific program (control flow), compiling it against a corresponding backend, such as LLVM, to execute it just-in-time.

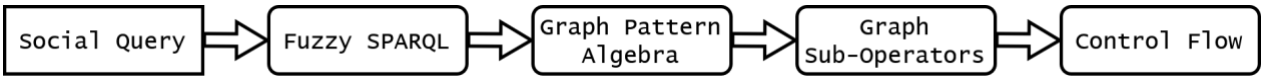


Figure 2: Prospective query processing pipeline for formalized social queries

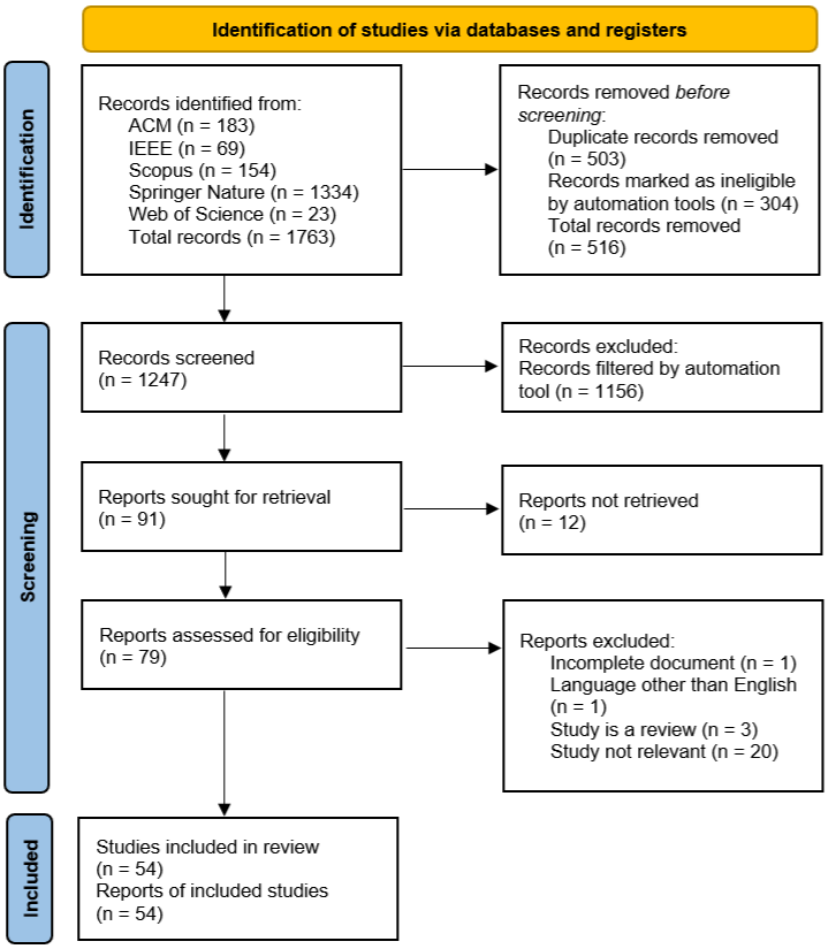
4. Conclusion

The authors contribute a systematic literature review on fuzzy techniques applied in social robotics following the PRISMA2020 scheme. The findings indicate that methods based on fuzzy logic are highly suitable for modelling a large set of different social abilities, such as social-cognitive movement, human emotion and intention prediction, social-cognitive communication, or simulation and expression of artificial emotions through behavior programming. Appendix 2 contains the review results per individual work. Based on these findings, the authors defined design goals for a prospective social knowledge base consisting of a fuzzy semantic database and a query processor to answer formalized social queries. Finally, the authors provided foundational aspects of an implementation approach for future research, consisting of a Property Graph data structure and a just-in-time compiled query engine for efficient execution of social queries formalized in a fuzzy superset of SPARQL.

Acknowledgements

The research described in this contribution is funded by BFS (de: Bayerische Forschungsstiftung) within the funding initiative FORSocialRobots (AZ-1594-23).

Appendix



Appendix 1: PRISMA2020 flow chart

Appendix 2: Review works with related social robot abilities as any from Social-cognitive movement and positioning (SCM), Human emotion and intention recognition (HEI), Simulation and expression of artificial emotions (SEE), Social-cognitive communication based on domain-specific knowledge (SCC), Artistic co-creation (ACC), Person recognition (PR), Social-cognitive admittance control (SCA), and Collaborative assistance and division of labor (CAL).

Ref.	Social Ability	Fuzzy Techniques	Social Query
[46]	SEE	Fuzzy logic with linguistic variables for perceived human distance, velocity, and robot's perceived emotion	Based on recognition of a person blocking a robot, which emotional response should be perceived?
[33]	HEI, SCC	Fuzzy logic with linguistic variables for hand gestures; learning system to update probabilistic fuzzy rules	Based on a perceived hand gesture, was an actual command to initiate a service issued by the user?
[39]	SCC	Fuzzy logic with linguistic variables for distances to objects and user requested object movements	Based on users' object movement commands, how much should the robot move the object?
[26]	SCC	Fuzzy semantic database with memberships of vocabularies to social contexts using fuzzy sets	Based on the robot's observed social context, which semantic meaning does apply for a term?
[47]	HEI	Fuzzy neural network with adaptive network-based fuzzy inference system for emotion classification	Based on the robot's perceived human utterance, what is the emotional state?

[24]	SEE	Fuzzy logic with fuzzy rule interpolation to model the behavior of a dog robot missing its owner	Based on the current observed state, what is the dog robot's anxiety and perceived missing of the owner?
[15]	HEI, SCC	Fuzzy logic with linguistic variables describing social atmosphere; behavior adaptation w. fuzzy Q-learning	Based on the perceived communication atmosphere, how should the robot adapt its behavior?
[48]	SCM	Fuzzy robot controller with linguistic variables for motor direction and motion changing frequency	Based on the forces received from the user, was it a navigational cue towards the dog robot?
[11]	HEI	Two-Layer Support Vector Regression model using Fuzzy C-Means for preliminary input data clustering	Based on the emotional state, age, gender and nationality, what drink could a bar service robot offer?
[49]	HEI	Fuzzy classifier: Supervised and Dynamic Fuzzy Adaptive System ART-based	Based on physical parameters, how should the rehabilitation robot adjust rehab. therapy difficulty?
[40]	SCC	Mamdani type fuzzy inference system for robot movement commands to determine the distance to travel.	Based on user commands and varying free space, how much must the robot move in its environment?
[50]	HEI	Neural network for posture classification using trapezoidal membership functions labelling pose features	Based on an observed human pose and facial expression, what is the emotional state?
[51]	HEI, SEE	Fuzzy logic with linguistic variables of pleasure and arousal intensities per emotion vector dimension	Based on the robot's observation, what is the human emotion intensity and the robot's facial response?
[22]	HEI, SEE	Fuzzy classifier for human emotion using Fuzzy C-Means on input; fuzzy rules for robot emotion states	Based on a human's face, what is the human's emotional state and resulting robot emotional state?
[36]	HEI, SEE, SCC	Type-I/II fuzzy logic with linguistic variables for input (pressure signals) and output (robot body gesture)	Based on the human interaction signals perceived, how should the therapy robot react to social stimuli?
[34]	HEI, SEE	Fuzzy logic with linguistic variables for human pose-extracted emotions; state machine for robot emotions	Based on observed human pose values, what is the current human and resulting robot emotional state?
[37]	HEI, SEE, SCC	Type-I/II fuzzy logic with linguistic variables for time and score to adjust a play activity's difficulty setting	Based on score and time in a playful activity with the robot, what is the game's adjusted difficulty?
[14]	HEI	Two-Layer SVR with Fuzzy C-Means for human emotion; Fuzzy Friend Q-Learning for behavior adaptation	Based on the emotional state, age, gender and nationality, what drink could a bar service robot offer?
[16]	HEI	Three-Layer SVR with Fuzzy C-Means for human emotion; weighted fusion model for human intention	Based on the emotional state, age, gender and nationality, what drink could a bar service robot offer?
[35]	HEI, SEE	Fuzzy classifier with C-Means for human emotion; fuzzy rules describing state machine of robot emotion	Based on a person's face, what is the current human and simulated robot emotional state?
[12]	HEI	Takagi-Sugeno fuzzy inference with linguistic variables for intention degree and supplied water quantity	Based on environmental and per-human observations, how should the service robot supply water?
[23]	SEE	Fuzzy logic with linguistic variables for environment, observed human and simulated robot emotional state	Based on environmental and per-human observations, what is the robot's simulated emotional state?
[41]	SCM	Fuzzy logic with linguistic variables for sensor detections and obstacle positions in social environments	Based on the robot's sensor detections, what is the approximate position of static or moving obstacles?
[13]	SCM, HEI	Takagi-Sugeno fuzzy inference with linguistic variables for approaching direction, distance and velocity	Based on an observed human's movement intention, how should the robot adjust its trajectory?
[52]	SCM, HEI, SEE, SCC	Fuzzy controllers for robot movement; fuzzy inference for human emotion deduction from input dialogue; fuzzy logic with linguistic variables for observed human and simulated robot emotional state	Based on the robot's sensor detections, what is the approximate position of obstacles? During communication, what is the assumed human emotional state and resulting robot emotional state?
[17]	HEI	Two-Layer Fuzzy Multiple Random Forest for speech emotion recognition with Fuzzy C-Means	Based on the perceived voice features, what is the human's emotional state?

[27]	HEI, SCC	Fuzzy logic with linguistic variables for user information, human emotion and robot utterance strategy	Based on given user information and human emotion, how should the robot convey compliments?
[8]	SCM	Fuzzy logic with linguistic variables for human proxemic zone, gender, familiarity and distance to robot	Based on gender, relative distance and familiarity, which proxemic zone may the robot enter?
[28]	ACC	Encoding of domain-specific musician heuristic knowledge in a fuzzy rule base	Based on the perceived properties of human improvised music, how should a robot dance to it?
[20]	HEI	Fuzzy DNN w. Sparse Autoencoder; Fuzzy Weighted Average based Fusion Model for human intention	Based on the emotional state, age, gender and nationality, what drink could a bar service robot offer?
[21]	HEI	Two-Layer SVR model for emotional intention; Takagi-Sugeno inference to deduce human intention state	Based on the emotional state, age, gender and nationality, what drink could a bar service robot offer?
[53]	SCM	Fuzzy logic with linguistic variables for obstacle distance, position, risk value and robot speed	Based on distance to human obstacles, the direction and risk, how should the robot adjust its speed?
[29]	PR	Mamdani type fuzzy inference system with linguistic variables signal-to-noise-ratio and utterance length	Based on the range of perceived utterance observations, who is the person speaking to the robot?
[54]	HEI	Interval Type-2 Recurrent Wavelet Fuzzy Brain Emotion Learning Network Model	Based on human facial expression and utterance, what is the human's emotional state?
[55]	SCM, HEI	Takagi-Sugeno fuzzy inference with linguistic variables for navigating agents' position and orientation	Based on a human's observed movement intention, how should the robot adjust its trajectory?
[30]	SCA	Fuzzy admittance controller with hedge algebra for external and transmitted force, and kinematic velocity	Based on the robot's experienced external force and linear velocity, how fast can the end-effector move?
[42]	SCM	Fuzzy social force model with linguistic variables for distance, linear velocity and interference behavior	Based on human distance, object velocity and behavioral strategy, how fast should the robot move?
[56]	HEI	Fuzzy neural network with linguistic variables for human distance, orientation and engagement intent	Based on the robot's observed human positioning and movement, does the person intend to interact?
[18]	HEI	Broad-Deep Network-Based Fuzzy Emotional Inference with fuzzy inference model to fuse observations	Based on observed facial expressions, behavior and personal information, what is the intention state?
[9]	SCM, SCC	Fuzzy Proxemic Evaluation Model with linguistic variables for robot movement and user feedback	Based on human pose, movement, and user feedback, what is the required interpersonal distance?
[32]	CAL	Mamdani fuzzy inference with linguistic variables for task ergonomics, safety and technical feasibility	Based on human ergonomics, safety, technical feasibility, should the robot assist a collaborative task?
[44]	HEI, SCC	Fuzzy logic to give meaning to social cues; fuzzy behavior rules for adaptation; learning of memberships	Based on the observed social cues and context, how should the robot adjust its behavior?
[57]	HEI	Fuzzy domain ontology based on a Fuzzy OWL extension for image feature annotation of panic levels	Based on the annotations on shared images in a social network, what is the human's emotional state?
[19]	HEI	Broad-Deep Network-Based Fuzzy Emotional Inference with fuzzy inference fusion model	Based on facial expression, behavior and personal information, what is the robot's teaching validity?
[58]	HEI	Fuzzy Transformer Model with three pipelines taking video, audio and textual inputs	Based on the observed video, audio and textual social data, what emotional nuances are perceived?
[25]	SEE	Ethological model based on fuzzy rules to determine animal aggression behavior and state transfers	Based on an animal robot's observations of familiarity and foes' distances, what is the behavior state?
[38]	HEI, SCC	Fuzzy logic with linguistic variables of observed emotional triggers, states and their changes	Based on speech extracted social cues, how does the robot's observed social situation change?

[10]	PR	Takagi-Sugeno fuzzy inference to fuse multiple recognition results based on a set of fuzzy rules	Based on face and body features, who is the person interacting with the robot?
[31]	SCA	Variable Operator Fuzzy Neural Network based on Takagi-Sugeno inference for admittance control	Based on the perceived external torque, how should the rehabilitation robot adjust its virtual damping?
[59]	HEI	Fuzzy CNN using fuzzy inference with exemplary linguistic variables of volume, tone, and human emotion	Based on the extracted human features a robot observes, what is the human's emotional state?
[60]	HEI, SCC	Fuzzy rules to supervise models for user speech analysis, facial expression and LLM text classification	Based on user facial emotion, speech sentiment and the robot's LLM utterance, how is user satisfaction?
[43]	SCM	Mamdani fuzzy inference with linguistic variables for object distances, turn intensity and robot speed	Based on the observed distance and relative velocity to objects, how should the robot adjust its speed?
[61]	SEE	Elliptic Intuitionistic Fuzzy Sets, a fuzzy theory utilizing two degrees of truth (falsehood and truth)	Based on the weighted input criteria, how strong should the robot smile?
[62]	HEI, SEE	2-Stage Fuzzy inference architecture to determine a robot's facial expression based on user facial expression	Based on facial observations, what is the human arousal, valence and resulting robot expression?

References

[1] C. Breazeal, K. Dautenhahn and T. Kanda, "Social Robotics," in *Springer Handbook on Robotics*, Berlin, Springer, 2016, pp. 1935-1972.

[2] H. Mahdi, S. A. Akgun, S. Saleh and K. Dautenhahn, "A survey on the design and evolution of social robots — Past, present and future," *Robotics and Autonomous Systems,* vol. 156, p. 104193, 2022.

[3] E. Durkheim, "What is a Social Fact?," in *Contemporary Social Theory*, London, Palgrave, 1982, pp. 50-59.

[4] W. Kritzinger, M. Karner, G. Traar, J. Henjes and W. Sihn, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine,* vol. 51, no. 11, p. 1016–1022, 2018.

[5] P. Hitzler, M. Krotzsch and S. Rudolph, Foundations of Semantic Web Technologies, New York: CRC Press Taylor & Francis Group, 2009.

[6] L. A. Zadeh, "Fuzzy logic," *Computer,* vol. 21, no. 4, pp. 83-93, 1988.

[7] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting and D. Moher, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ,* vol. 372, p. 71, 2021.

[8] P. Patompak, S. Jeong, I. Nilkhamhang and N. Y. Chong, "Learning Proxemics for Personalized Human–Robot Social Interaction," *International Journal of Social Robotics,* vol. 12, no. 1, pp. 267-280, 2019.

[9] S. M. B. P. Samarakoon, M. A. V. J. Muthugala, A. G. B. P. Jayasekara and M. R. Elara, "Adapting approaching proxemics of a service robot based on physical user behavior and user feedback," *User Modeling and User-Adapted Interaction,* vol. 33, no. 2, pp. 195-220, 2023.

[10] C. Z. Siow, W. Wang, T. Obo and N. Kubota, "Continuous Person Recognition on Balloon Robot Using TSK Fuzzy Classifier with Face and Body Regions for Social Distance Purposes," in *2024 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Yokohama, Japan, 2024.

[11] L.-F. Chen, Z.-T. Liu, M. Wu, M. Ding, F.-Y. Dong and K. Hirota, "Emotion-Age-Gender-Nationality Based Intention Understanding in Human–Robot Interaction Using Two-Layer Fuzzy Support Vector Regression," *International Journal of Social Robotics,* vol. 7, no. 5, pp. 709-729, 2015.

[12] M. Hao, W. Cao and Z. Liu, "A Demand Analysis Method Based on Takagi-Sugeno Fuzzy Model for Drinking Service," in *2018 37th Chinese Control Conference (CCC)*, Wuhan, China, 2018.

[13] R. Palm, R. Chadalavada and A. J. Lilienthal, "Fuzzy Modeling, Control and Prediction in Human-Robot Systems," in *International Joint Conference on Computational Intelligence*, Vienna, Austria, 2019.

[14] L. Chen, M. Wu, M. Zhou, J. She, F. Dong and K. Hirota, "Information-Driven Multirobot Behavior Adaptation to Emotional Intention in Human–Robot Interaction," *IEEE Transactions on Cognitive and Developmental Systems,* vol. 10, no. 3, pp. 647-658, 2017.

[15] L.-F. Chen, Z.-T. Liu, M. Wu, F.-Y. Dong, Y. Yamazaki and K. Hirota, "Multi-robot behavior adaptation to local and global communication atmosphere in humans-robots interaction," *Journal on Multimodal User Interfaces,* vol. 8, no. 3, pp. 289-303, 2014.

[16] L. Chen, M. Zhou, M. Wu, J. She, Z. Liu, F. Dong and K. Hirota, "Three-Layer Weighted Fuzzy Support Vector Regression for Emotional Intention Understanding in Human–Robot Interaction," *IEEE Transactions on Fuzzy Systems,* vol. 25, no. 5, pp. 2524-2538, 2018.

[17] L. Chen, W. Su, Y. Feng, M. Wu, J. She and K. Hirota, "Two-layer fuzzy multiple random forest for speech emotion recognition in human-robot interaction," *Information Sciences,* vol. 509, no. 1, pp. 150-163, 2020.

[18] M. Li, L. Chen, M. Wu, M. Ding, K. Hirota and W. Pedrycz, "Broad-Deep Network-Based Fuzzy Emotional Inference Model with Personal Information for Intention Understanding," *IFAC-PapersOnLine,* vol. 56, no. 2, pp. 7071-7076, 2023.

[19] M. Li, L. Chen, M. Wu and K. Hirota, "A broad-deep fusion network-based fuzzy emotional intention inference model for teaching validity evaluation," *Information Sciences,* vol. 654, no. 1, p. 119837, 2024.

[20] L. Chen, W. Su, M. Wu, W. Pedrycz and K. Hirota, "A Fuzzy Deep Neural Network With Sparse Autoencoder for Emotional Intention Understanding in Human–Robot Interaction," *IEEE Transactions on Fuzzy Systems,* vol. 28, no. 7, pp. 1252-1264, 2020.

[21] L. Chen, M. Wu, M. Zhou, Z. Liu, J. She and K. Hirota, "Dynamic Emotion Understanding in Human–Robot Interaction Based on Two-Layer Fuzzy SVR-TS Model," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 50, no. 2, pp. 490-501, 2020.

[22] A. Meghdari, M. Alemi, A. G. Pour and A. Taheri, "Spontaneous Human-Robot Emotional Interaction Through Facial Expressions," in *International Conference on Social Robotics*, Kansas City, USA, 2016.

[23] L. Rincon, E. Coronado, H. Hendra, J. Phan, Z. Zainalkefli and G. Venture, "Expressive states with a robot arm using adaptive fuzzy and robust predictive controllers," in *2018 3rd International Conference on Control and Robotics Engineering (ICCRE)*, Nagoya, Japan, 2018.

[24] D. Vincze, S. Kovács, M. Niitsuma, H. Hashimoto, P. Korondi, M. Gácsi and Á. Miklósi, "Ethologically Inspired Human-Robot Interaction Interfaces," in *Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments*, New York, NY, USA, 2012.

[25] M. A. Lone, O. M. Khanday and S. Kovács, "Implementation Guidelines for Ethologically Inspired Fuzzy Behaviour-Based Systems," *Infocommunications Journal,* vol. 16, no. 3, pp. 43-56, 2024.

[26] S. Jusoh and H. M. Alfawareh, "Handling Ambiguous Entities in Human-Robot Instructions Using Linguistic and Context Knowledge," in *Digital Information and Communication Technology and Its Applications*, Berlin, Heidelberg, 2011.

[27] M. Iwashita and D. Katagami, "Psychological Effects of Compliment Expressions by Communication Robots on Humans," in *2020 International Joint Conference on Neural Networks (IJCNN)*, Glasgow, UK, 2020.

[28] O. Thörn, P. Knudsen and A. Saffiotti, "Human-Robot Artistic Co-Creation: a Study in Improvised Robot Dance," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Naples, Italy, 2020.

[29] M. Al-Qaderi, E. Lahamer and A. Rad, "A Two-Level Speaker Identification System via Fusion of Heterogeneous Classifiers and Complementary Feature Cooperation," *Sensor,* vol. 21, 2021.

[30] N.-V. Toan, P.-B. Khoi and S.-Y. Yi, "A MLP-Hedge-Algebras Admittance Controller for Physical Human–Robot Interaction," *Applied Sciences,* vol. 11, no. 12, p. 5459, 2021.

[31] J. Li, Y. Zhou, S. Zuo and M. Dong, "Dynamically modulated robot compliance via online fuzzy neural networks for individualized ankle rehabilitation," *Applied Intelligence,* vol. 55, no. 6, p. 411, 2025.

[32] A. Alessio, K. Aliev and D. Antonelli, "Multicriteria task classification in human-robot collaborative assembly through fuzzy inference," *Journal of Intelligent Manufacturing,* vol. 35, no. 5, pp. 1909-1927, 2024.

[33] Z. Z. Bien, H.-E. Lee, J.-H. Do, Y.-H. Kim, K.-H. Park and S.-E. Yang, "Intelligent Interaction For Human-Friendly Service Robot in Smart House Environment," *International Journal of Computational Intelligence Systems,* vol. 1, no. 1, pp. 77-93, 2008.

[34] T. Zhu, Q. Zhao and J. Xiong, "A Body Emotion-Based Human-Robot Interaction," in *International Conference on Computer Vision Systems*, Shenzhen, China, 2017.

[35] A. Ghorbandaei Pour, A. Taheri, M. Alemi and A. Meghdari, "Human–Robot Facial Expression Reciprocal Interaction Platform: Case Studies on Children with Autism," *International Journal of Social Robotics,* vol. 10, no. 2, pp. 179-198, 2018.

[36] P. Ponce, A. Molina und D. Grammatikou, „Design based on fuzzy signal detection theory for a semi-autonomous assisting robot in children autism therapy," *Computers in Human Behavior,* Bd. 55, Nr. 1, pp. 28-42, 2016.

[37] P. Ponce, A. Molina, D. Grammatikou and O. Matar, "Fuzzy Logic Type 1 and 2 for Social Robots and Apps for Children with Autism," in *2017 Sixteenth Mexican International Conference on Artificial Intelligence*, Mexico City, Mexico, 2017.

[38] D. Vincze and M. Niitsuma, "A Human Speech-based Behavior Model for Social Robots based on Parent-Child Interaction Therapy," in *2024 IEEE/SICE International Symposium on System Integration (SII)*, Ha Long, Vietnam, 2024.

[39] A. G. B. P. Jayasekara, K. Watanabe and K. Izumi, "Understanding user commands by evaluating fuzzy linguistic information based on visual attention," *Artificial Life and Robotics,* vol. 14, no. 1, pp. 48-52, 2009.

[40] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "Interpreting fuzzy linguistic information in user commands by analyzing movement restrictions in the surrounding environment," in *2015 Moratuwa Engineering Research Conference (MERCon)*, Moratuwa, Sri Lanka, 2015.

[41] S. D. N. Tanha, S. F. Dehkordi und A. H. Korayem, „Control a mobile robot in Social environments by considering human as a moving obstacle," in *Proceedings of the 6th RSI International Conference on Robotics and Mechatronics (IcRoM 2018)*, Tehran, Iran, 2018.

[42] X. Gao, Q. Zhang, M. Li, B. Lan, X. Fu and J. Li, "Mobile Robot Combination Autonomous Behavior Strategy to Inspect Hazardous Gases in Relatively Narrow Man–Machine Environment," *Chinese Journal of Mechanical Engineering,* vol. 35, no. 1, p. 135, 2022.

[43] L. C. Sousa, Y. M. R. Silva, V. B. Schettino, T. M. B. Santos, A. R. L. Zachi, J. A. Gouvêa and M. F. Pinto, "Obstacle Avoidance Technique for Mobile Robots at Autonomous Human-Robot Collaborative Warehouse Environments," *Sensors,* vol. 25, no. 8, p. 2387, 2025.

[44] D. Dell'Anna and A. Jamshidnejad, "SONAR: An Adaptive Control Architecture for Social Norm Aware Robots," *International Journal of Social Robotics,* vol. 16, no. 9, p. 1969–2000, 2024.

[45] M. Jungmair, A. Kohn and J. Giceva, "Designing an open framework for query optimization and compilation," *Proc. VLDB Endow.,* vol. 15, no. 11, p. 2389–2401, 2022.

[46] H. Mobahi and S. Ansari, "Fuzzy perception, emotion and expression for interactive robots," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483)*, Washington, DC, USA, 2003.

[47] N. Kamaruddin, A. Wahab and C. Quek, "Cultural dependency analysis for understanding speech emotion," *Expert Systems with Applications,* vol. 39, no. 5, pp. 5115-5133, 2012.

[48] Y. Wei and M. Lee, "A Guide-dog Robot System Research for the Visually Impaired," in *2014 IEEE International Conference on Industrial Technology (ICIT)*, Busan, Korea, 2014.

[49] L. D. Lledó, F. J. Badesa, M. Almonacid, J. M. Cano-Izquierdo, J. M. Sabater-Navarro, E. Fernández and N. Garcia-Aracil, "Supervised and Dynamic Neuro-Fuzzy Systems to Classify Physiological Responses in Robot-Assisted Neurorehabilitation," *PLOS ONE,* 22 May 2015.

[50] T. Obo, L. C. Kiong and N. Kubota, "Human behavior modeling for multimodal interaction with robot partner," in *2015 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*, Nagoya, Japan, 2015.

[51] Z. Rasool, N. Masuyama, M. N. Islam and C. K. Loo, "Empathic Interaction Using the Computational Emotion Model," in *2015 IEEE Symposium Series on Computational Intelligence*, Cape Town, South Africa, 2015.

[52] B. P. E. A. Vásquez and F. Matía, "A tour-guide robot: Moving towards interaction with humans," *Engineering Applications of Artificial Intelligence,* vol. 88, no. 1, p. 103356, 2020.

[53] A. Terra, H. Riaz, K. Raizer, A. Hata and R. Inam, "Safety vs. Efficiency: AI-Based Risk Mitigation in Collaborative Robotics," in *2020 6th International Conference on Control, Automation and Robotics (ICCAR)*, Singapore, 2020.

[54] D. Li and X. Chen, "Emotion Recognition Based on Type-2 Recurrent Wavelet Fuzzy Brain Emotion Learning Network Model," *Mathematical Problems in Engineering,* vol. 2021, p. 9991531, 2021.

[55] R. Palm and A. J. Lilienthal, "Fuzzy Geometric Approach to Collision Estimation Under Gaussian Noise in Human-Robot Interaction," in *International Joint Conference on Computational Intelligence*, Valletta, Malta, 2021.

[56] J. Bi, F. Hu, Y. Wang, M. Luo and M. He, "Human engagement intention intensity recognition method based on two states fusion fuzzy inference system," *Intelligent Service Robotics,* vol. 16, no. 3, pp. 307-322, 2023.

[57] O. Lazzez, T. M. Hamdani, H. Chabchoub, K. Ouhada and A. M. Alimi, "Fuzzy deep visual panic detection," *Social Network Analysis and Mining,* vol. 14, no. 1, p. 180, 2024.

[58] J. Liu, M. C. Ang, J. K. Chaw, K. W. Ng and A.-L. Kor, "Personalized emotion analysis based on fuzzy multi-modal transformer model," *Applied Intelligence,* vol. 55, no. 3, p. 227, 2024.

[59] T. Meena, V. A. Devi, A. Nagpal, P. Srivasatava, S. Ziara and V. Sailaja, "Adaptive Emotion Recognition Framework Leveraging Lightweight Convolutional Networks and Fuzzy Logic for Enhanced Interpretability and Efficiency in Low-Resource Environments," in *2025 International Conference on Intelligent Control, Computing and Communications (IC3)*, Mathura, India, 2025.

[60] K. A. Shaik, S. Xie, F. Cruz and E. B. Sandoval, "A Fuzzy Supervisory Framework for Real-Time Optimization of Robot Output and LLM Performance in HRI," in *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Melbourne, Australia, 2025.

[61] V. Traneva, S. Tranev and V. Tododorov, "Modeling Humanoid Robot Mouth Expressions Reflecting Emotional States Via Elliptic Intuitionistic Fuzzy Sets," *Journal of Theoretical and Applied Mechanics,* vol. 55, no. Special Issue, pp. 35-51, 2025.

[62] N. M. Trieu and N. T. Thinh, "Enhancing Emotional Expressiveness in Biomechanics Robotic Head: A Novel Fuzzy Approach for Robotic Facial Skin's Actuators," *Computer Modeling in Engineering & Sciences,* vol. 143, no. 1, pp. 477-498, 2025.

Biography

Raoul Zebisch (*1999) holds a position as a research associate at the Chair of Digital Manufacturing at the University of Augsburg since 2023, researching the virtual validation of social robots through process-oriented simulations with semantic knowledge bases.

Johannes Schilp (*1976) has been Chair of Digital Manufacturing since 2015. He is also head of the processing department at the Fraunhofer Research Institution for Casting, Composite and Processing Technology and a member of scientific networks like WG MHI and automation research groups (VDMA).

8th Conference on Production Systems and Logistics

Implementation and Validation of Elastic Transformation Engine for Asset Administration Shells

Benjamin Goetz[1], Matthias Schneider[1], Daniel Schel[1], Thomas Bauernhansl[1,2]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstrasse 12, 70569 Stuttgart, Germany*
[2]*Institute of Industrial Manufacturing and Management, University of Stuttgart, Nobelstraße 12, 70569 Stuttgart, Germany*

Abstract

The introduction of the Asset Administration Shell as an implementation concept for the Digital Twin of industrial assets, represents a crucial step toward achieving holistic interoperability between industrial hardware and software. However, the standardization process for data model structures, known as submodel templates, can be time-consuming. As a result, implementing and updating the Asset Administration Shell on the shop floor presents new challenges. One solution to address these challenges is the development of an Asset Administration Shell transformation engine for submodels. This engine can convert the structure of submodels in cases of submodel template revisions or modify the values of submodels for data preprocessing. The objective of this paper is to describe and implement a reference implementation of this concept of a transformation engine and to validate its implementation with real-world use cases. The implementation emphasizes a service-oriented software design, which allows the engine to be flexible and scalable according to workload demands. This approach ensures that transformed submodels can be provided based on the specific requirements of the shop floor. Validation consists of two test cases: the first simulates the live conversion of a machine's submodel that contains sensor data; the second simulates the data preprocessing of a time series of historical machine sensor data. Key performance indicators are identified to evaluate the engine's behavior during these tests and collect relevant data. The results are analyzed with a focus on whether the service-oriented design effectively manages the workload for the test cases and how the submodel transformation process can be improved in future updates of the implementation.

Keywords

Asset Administration Shell; Transformation; Cloud; Service-Oriented Architecture; Scaling

1. Introduction

As the Asset Administration Shell (AAS) is actively being developed towards becoming the standard for factory digital twins, it is increasingly important to handle the data provided by shells and submodels at the volumes and frequencies required by modern production processes [1,2]. Data handling involves, on the one hand, processing the incoming data and, on the other hand, transforming the data structure to meet company- or domain-specific conventions or standards. Even once introduced, conventions and standards may evolve and require transformation processes. This challenges factory operators when trying to provide a real-time, standardized digital representation of their factories. To address this challenge, this publication examines whether a horizontally scalable transformation engine can transform AAS submodels without compromising data integrity.

DOI: https://doi.org/10.15488/20971
ISSN: 2701-6277

The paper gives a brief introduction into the theoretical background of the topic in section 2. Section 3 describes the software design of an existing proposal of an AAS transformation engine [3]. Section 4 contains a description of implementation details relevant to understanding and discussing the later performance tests. Section 5 introduces two test scenarios and a test environment to test the implementation at scale. The tests focus on the effect of scaling the transformation engine's services on increasing the transformation performance. Finally, Section 6 summarizes the results and gives an outlook on improving the implementation.

2. Theoretical Background

To enhance the quality of production processes continuously, it is essential to collect machine data and assess the quality of production steps. Access to production data, combined with modern data analytics, enables companies to identify patterns and anomalies in the data collected throughout the production process. This information, along with product quality indicators, aids in identifying sources of errors and establishing feedback loops for ongoing improvements in production [4–6].

As modern production supply chains involve multiple companies, it is imperative to access data that spans organizational boundaries. The introduction of the Asset Administration Shell (AAS) as a manifestation of the Digital Twin facilitates a centralized access point to all product data, even when this data is divided into various parts – termed submodels in the context of AAS – and provided by different entities. An example of this is the production of electrolyzers, which entails numerous production steps carried out by specialized small and medium-sized enterprises. [7–9].

When setting up data processing pipelines, there is a consistent need to translate between the provided and expected data structures, as well as to process data to eliminate errors and noise [10–13]. In the context of the AAS, this means that a transformation engine must be capable of transforming the structure as well as the values of a submodel. Depending on the feedback loop's design, especially when it aims to control production processes directly, the solution must effectively handle data at high frequencies and large data sets, providing real-time feedback. To meet these performance requirements, the proposed solution will adopt software design principles utilized in cloud computing [3].

Cloud systems enable scalability across hardware, virtualization, and software layers. This empowers operators to run software elastically, allowing it to scale – ideally automatically – based on current workloads to minimize cloud resource usage while ensuring a defined level of service [14–16]. While cloud providers manage the scalability of hardware and virtualization layers, software architects must design their applications according to scalable design patterns. One such pattern is the use of microservices, which stems from service-oriented architectures (SoA; see Section 3 for more Details) and breaks applications into smaller, independent, and reusable services, enhancing flexibility and deployment speed. To process user requests, an initial request is divided into a sequence of smaller requests processed by microservices and returned as aggregated response to the user. The deployment of microservices employs container technology, simplifying the task for service providers to maintain the desired number of service replicas for high availability or load balancing.

For load balancing, a load balancer service is positioned in front of a service interface to distribute incoming requests among service replicas, enabling automatic scaling routines by monitoring service load and defining scaling rules. Load balancers may use static or dynamic strategies to distribute the load. Static load balancing relies on initial assumptions and predefined heuristics, while dynamic adapts to changing runtime conditions but create overhead [17–20].

An alternative to traditional load balancing is the use of a centralized queue. In this architecture, incoming requests are buffered within the queue and subsequently retrieved by consuming services capable of processing them. This mechanism shifts the system logic from a push-based to a pull-based principle. By

allowing consumers to request tasks at their own capacity, the approach inherently prevents resource exhaustion and eliminates the need for direct, stateful communication channels between provider and consumer [21,22]. The software design of the proposed transformation engine is based on this theoretical background and is described in the following section.

3. Software Design

The design of the transformation engine follows the principle of a service-oriented architecture (SoA) (see Figure 1) [23–25]. This design pattern divides the transformation engine into smaller microservices, each handling a specific sub-task of the overall task. To enable microservices to interact with each other, each service provides an interface for communicating with other services. This approach enables the system operator to scale the number of instances of a service type horizontally, thereby balancing the workload of the service type among the running instances. As depicted in Figure 1, the software design identifies two main service types for the transformation of submodels:

- Listener: listening to changes of submodels (step 1), creating transformation jobs (step 2), and sending the jobs to the in-memory cache (step 3) of the transformation engine, which acts as a job queue.
- Executor: executing submodel transformations based on the jobs cached in the queue. A transformation is based on a source submodel and transformation rules defined by the engine user (steps 3, 4, and 5).

Both service types are implemented in that way to run in parallel with multiple instances and not interfere with each other.

The communication between listeners and executors is handled via a job queue (within the cache of the engine). Listeners create and store transformation jobs (see Section 4.1) in the queue, while executors pull jobs from the queue and process them. The job queue serves as a buffer for jobs in case the listener instances

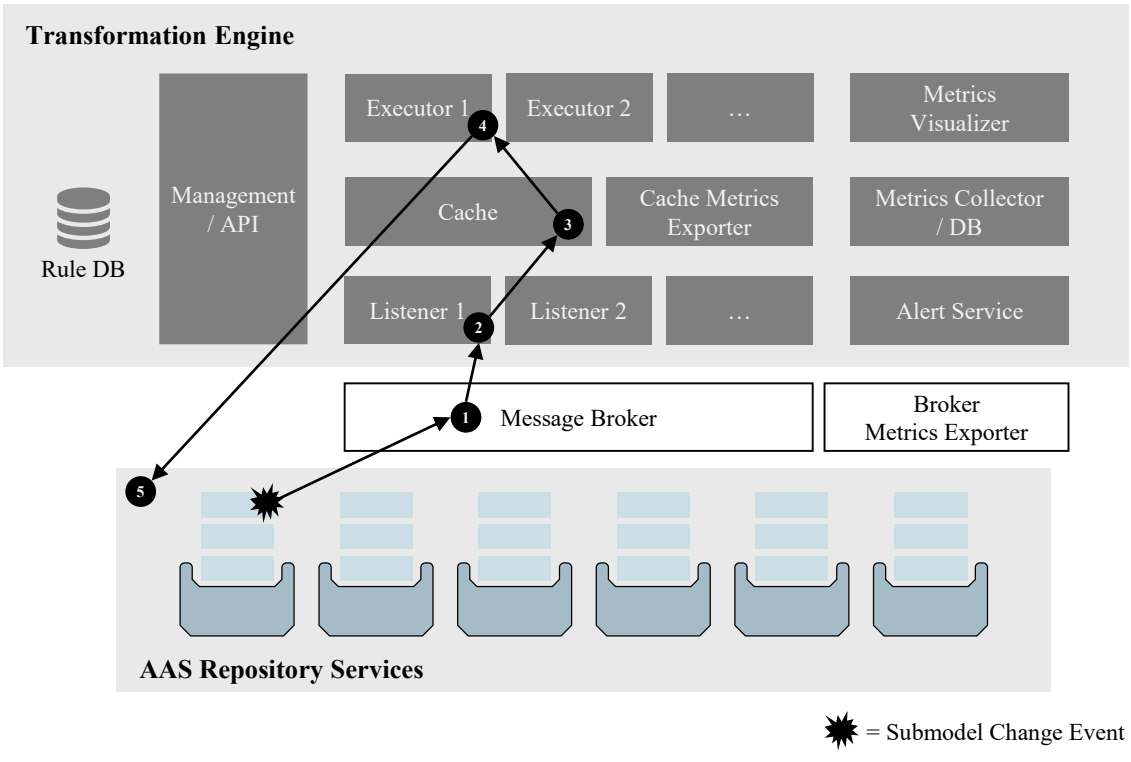


Figure 1: Software Design and Transformation Workflow

create jobs faster than the executor instances can process them. The result of processing a transformation job is a new or updated submodel that is sent to and saved by an external submodel repository by the executor.

Whether a submodel change causes a listener to create a transformation job depends on the list of transformation rules defined by the user. To manage these rules, the transformation engine provides a Management service that offers an API for creating, updating, and deleting rules. Every change to a rule is forwarded to the listeners and executors, allowing the services to adjust their behavior according to the updates of the transformation rules.

To measure the load on listeners and executors, the architecture is extended with a metrics database, a metrics visualizer, and an alert service. The alert service informs the runtime environment of the service type that needs to be scaled up or down to adapt to the current load on the service instances and optimize resource utilization. Metrics exporters for the message broker and the cache provide the relevant metrics, which are persisted in the metrics database (see Section 4.2 for details about metrics and load-balancing).

4. Implementation

4.1 Data Model of Transformation Job

A transformation job is a shared data model for communication between listeners and executors. Listeners create jobs and enqueue them, executors consume them. Each job must include all necessary information for the transformation, specifically the action type: EXECUTE when a source submodel is created/changed (run the transformation rule) and DELETE when the source submodel is deleted. The target submodel should also be deleted.

The source submodel is stored either by reference or by value in a job. By reference, only the submodel ID is stored; the executor loads the source submodel at start from an external submodel repository. Handling the source submodel by reference reduces job size and queue storage but risks processing the latest state instead of the one that actually triggered the job, potentially skipping transformations.

If every change must be processed, the source submodel can be stored by value in the job. The executor retrieves the source submodel from the job, rather than looking it up, which is more robust under frequent changes but increases demands on queue storage. At very high change rates, the engine may fall behind in fulfilling transformations, and the executor load may rise. Finally, each job also includes the target submodel ID required for distributed lock management (see Section 4.3).

The configuration of the engine affects storage, compute, and throughput, and should align with change rates of the source submodels and expectations, aiming either to reduce load or to capture every change, with clear trade-offs.

4.2 Load-Balancing and Metrics

Managing the transformation engine's load involves distributing work across service instances and collecting metrics to decide if scaling is needed. The engine can scale listener and executor services: the listener instance count aligns with incoming submodel changes from the message broker, and the executor count aligns with queued transformation jobs. Subsequent sections outline load-balancing and metric collection strategies, emphasizing service-specific metrics over basic runtime performance metrics (e.g., CPU or RAM utilization).

4.2.1 Listener

Open source implementations of AAS-hosting software, such as Eclipse AASX Server [26], Eclipse BaSyx [27], or FA³ST Service [28,29] utilize Message Queuing Telemetry Transport (MQTT) [30] for

communicating changes to submodels [31]. Therefore, for the listener, we are focusing on load-balancing mechanisms for messages sent via MQTT.

MQTT shared subscriptions allow multiple clients to join a shared group for a topic; each message goes to only one client, rotating via round robin, evenly distributing incoming submodel change notifications [32] (see Figure 2). However, MQTT's push message delivery mechanism can create high local cache demand on listener instances in case of heavy load. These local caches are not accessible to newly added listeners when scaling up, causing uneven load. To address this, broker messages are first pushed into the central cache of the transformation engine, then pulled by listeners depending on their individual load. Pushing the messages into the cache first removes per-listener caches and turns MQTT's push into a pull delivery mechanism that adapts to changing listener counts and per-instance load. To estimate load and scaling needs, the engine tracks n_t, the *total number of submodel changes*, and n_{np}, the *number of not-processed changes*. Because MQTT lacks per-topic metrics, the Management service counts n_t by subscribing to the same topics (outside the shared group), and n_{np} by reading the transformation engine's cache list of unprocessed change messages (exposed by the cache metrics exporter, see Section 3). Based on these metrics, the listener service is scaled accordingly.

4.2.2 Executor

As depicted in Figure 3, executor load-balancing utilizes two lists in the transformation engine cache: waiting jobs (pushed by listeners) and in-process jobs (currently handled by executors). Executors may be in an idle or working state and process only one job at a time. Idle executors poll the waiting list; the first to detect a new job moves it to the in-process list, loads a local copy, executes the job, then deletes it from the in-process list and resumes polling. This pull-based design spreads work evenly across executors, adapts to variable job durations, and prevents overload. Metrics are derived from these lists: the *number of waiting jobs* n_w, and the *number of in-process jobs* n_{ip}. The cache metrics exporter reads list lengths from the queue/cache and exposes them via an API for monitoring and scaling decisions (see Section 3). The *total count of executor services* n_e is utilized to calculate the *count of idle executors* by subtracting n_{ip} from n_e. Using these metrics, the overall executor load can be estimated, and informed scaling decisions can be made.

4.3 Distributed Lock Management

To ensure correct execution order and prevent conflicting parallel writes, the engine uses distributed lock management [33]. Executors place locks on target submodels at the start of a job; while locked, any queued jobs targeting the same submodel are blocked from being pulled or processed by other executors. The creating executor releases the lock after completion; if it cannot, each lock has a predefined time-to-live (TTL) and expires automatically, preventing locks to be stuck. Locks are stored in the transformation engine's cache, which enforces visibility to all executors and handles expiration. This lock management preserves job sequence and data integrity.

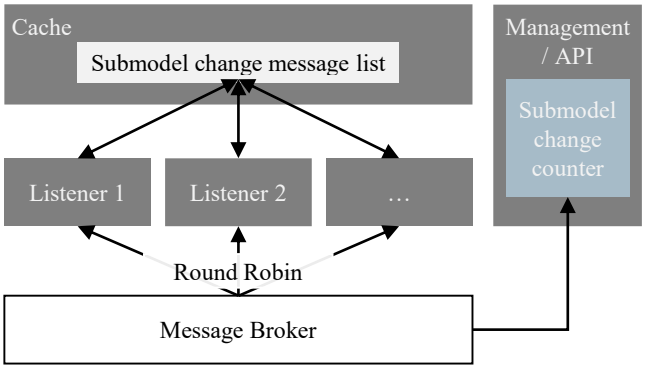


Figure 2: Load-balancing of listeners and measuring of metrics

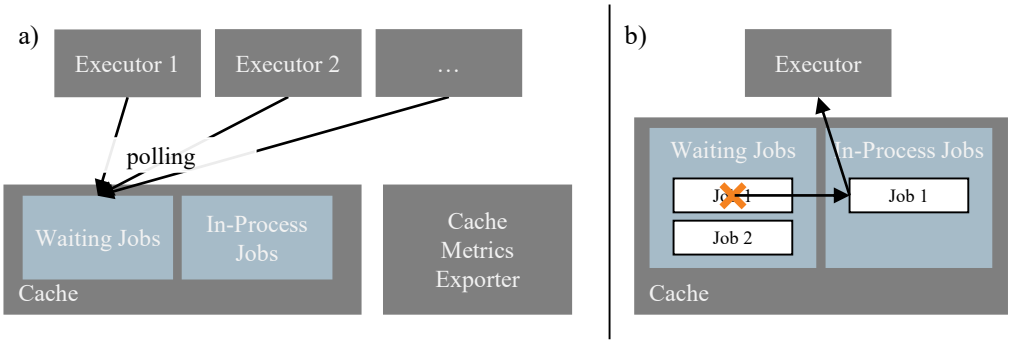


Figure 3: Executor load-balancing (a) and process of pulling job from queue (b)

5. Validation

The previous chapter described the details of the implementation and monitoring concept of the transformation engine. Based on this foundation, the service-oriented design of the engine is validated using real-world use cases. The validation shall demonstrate that the engine can meet its service level under load with the assistance of horizontal scaling of the listener and executor services.

5.1 Use Cases

The use cases for validation are machine-focused. The first case, named “machine data case”, consists of a machine publishing its live sensor data via a submodel based on the standardized submodel template of time series data [34]. The submodel contains one record for all sensors at present and updates the submodel at a high frequency as time progresses. The transformation used for validation renames the IDs of the sensors to simulate a version update of the submodel template, which required the renaming action.

The second use case, named “historic data case”, takes recordings of all sensors of a machine over a period of time as the input of the transformation. The recordings are also based on the submodel template of time series data (similar to the first, but with multiple records) and reflect data collected during one production step of a product. The transformation calculates the moving average of all sensor data and simulates the data preprocessing step of a machine learning lifecycle. While the first use case creates only one submodel that is updated continuously, the second one creates new submodels periodically, which remain unchanged after they have been created.

5.2 Test Environment and Setup

The transformation engine runs on a scalable Docker Swarm of 7 nodes, each with an Ubuntu 24.04.1 VM, featuring 8 CPU cores and 16 GB of RAM. All transformation engine’s services run as Docker containers. One node is dedicated to listeners and one to executors to prevent interference and ensure sufficient resources when scaling. Each listener and executor reserves 2 CPUs and 2 GB RAM.

Tests use two deployment variations to isolate scaling effects: when testing listeners, no executors run, and vice versa. Before each test, shells, submodels, and cache of transformation engine are cleared. A load is applied for 30 seconds using machine or historical data creators to simulate the machines. The number of creators is determined in pre-tests to ensure oversaturation of the listeners/executors. Afterward, service-specific KPIs are measured to assess the impact of scaling. KPIs are computed per second (KPI divided by test duration) for overall throughput, and per instance (KPI divided by instance count) to gauge load distribution. Tests run with 1, 2, and 3 instances, each repeated 5 times for statistical analysis.

An AAS test infrastructure is deployed via BaSyx’s off-the-shelf components in the swarm, including shell and submodel registry, and a combined shell/submodel repository (AAS Environment). Repositories send MQTT notifications on shell/submodel changes, which the transformation engine consumes as described in Section 3.

5.3 Test Results

5.3.1 Listener

To create a load on the listener instances, the machine data and historical data creators send submodels to the submodel repository. The repository sends change notifications to an MQTT broker that the listeners are listening to. The count of the creators has been set to 6.

The KPI for the listener tests is the number of transformation jobs created during the test run. After every test, the count of unprocessed MQTT messages is taken to ensure that the listeners have been oversaturated and have not been able to process the change notifications created during the test. This avoids negatively affecting the KPIs by not fully utilizing the instances.

Figure 4 summarizes the test results. The average number of created transformation Jobs per second is increasing with the increase in listener count for both test scenarios. This behavior meets the expectations of being able to handle more MQTT messages by scaling the listener instances up. The graphs of the average number of created transformation jobs per instance show a flat curve for both scenarios, while the KPIs of the machine data test increase from 1 to 2 listeners and decrease from 2 to 3 listeners. The KPIs of the historical data test show a slight decrease from 1 to 3 listeners. This means that all instances have been evenly utilized in all test scenarios.

5.3.2 Executor

In executor tests, data creators produce source submodels and transformation jobs because no listeners are running. The source submodel is included by value in each job (see the description of saving source submodels in job by reference/by value in Section 4.1), ensuring that every job triggers a transformation. Two creators are used, and the KPI is the number of transformations executed during the 30-second test. After each run, the number of remaining unprocessed jobs is counted to confirm that the executors were oversaturated and could not finish all jobs, preventing underutilization from skewing KPIs. Figure 5 summarizes the test results.

Historical data results mirror listener tests: the number of transformations per second rises with more executor instances, while per-instance throughput remains constant. Thus, scaling executors for historical data delivers the expected capacity gains.

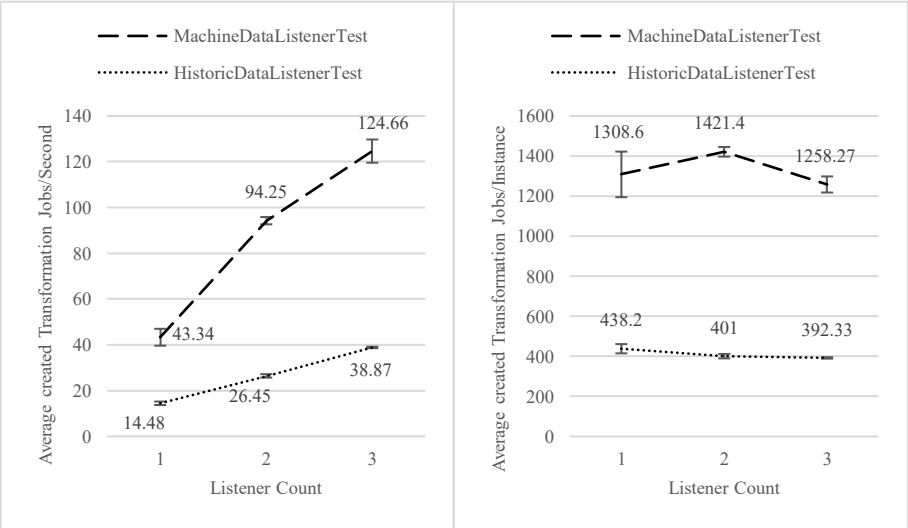


Figure 4: Listener test results with six data creators, five tests per listener count and a test duration of 30 seconds; left: average created transformation jobs per second over listener count; right: average created transformation jobs per instance over listener count

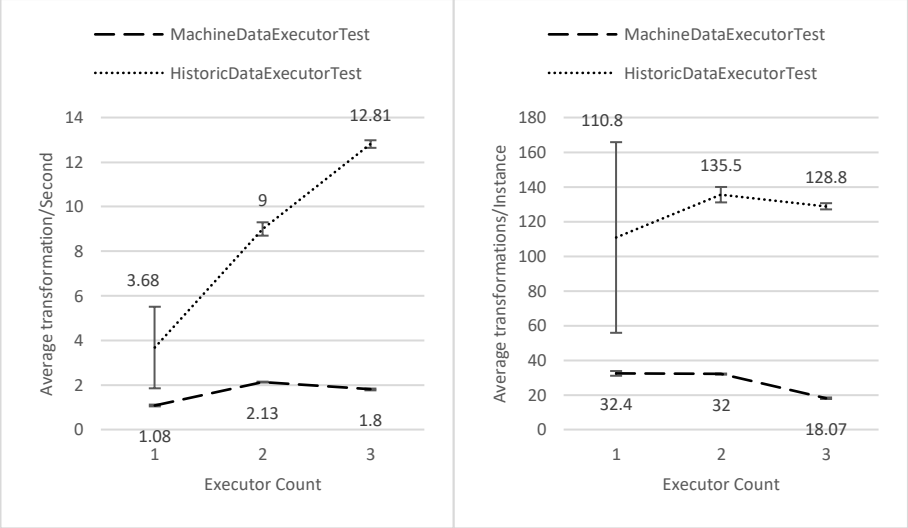


Figure 5: Executor test results with two data creators, five tests per executor count and a test duration of 30 seconds; left: average transformations per second over executor count; right: average transformations per instance over executor count

Machine data results differ: the number of transformations per second increases slightly from one to two executors but drops from two to three, contrary to expectations. Per-instance throughput is 32.4 and 32 for one and two instances, respectively, then falls to 18.07 with three, indicating even utilization at one and two instances but a decline at three. Adding a third instance yields no performance benefit and lowers load per instance. The cause is the distributed lock management (see Section 4.3): it blocks concurrent jobs targeting the same submodel. With two continuously updated submodels by the two machine data creators, only two of the three executors can work in parallel. The third waits for locks to be released.

6. Final Thoughts and Outlook

The main finding from test results indicate that scaling a service-oriented AAS transformation engine, as described in Section 4, does not meet the expectations for all scenarios and service types. Scaling the service types of the transformation engine only helps with submodels that are not updated or are updated infrequently, such as historic data. However, when a data source changes a submodel at a high frequency, scaling up the transformation engine's executor services does not meet the load. The integrated lock management, which ensures the data integrity of submodels, mitigates the performance impact of scaling. Therefore, the service-oriented submodel transformation concept needs adjustment to support high-frequency data sources, such as sensor data of shop floor machines.

One promising solution is to offer an alternative behavior for rapidly changing source submodels: instead of executing transformations and saving results every time a source submodel changes, the transformation would only occur when a target submodel is requested. This transformation-on-demand approach could eliminate the need for distributed lock management and reduce engine load when source submodel updates happen more frequently than requests for transformation results. Furthermore, this method may improve the effect of load balancing strategies while still ensuring data integrity.

Acknowledgements

This paper is based on the work in the research projects H2Giga-FRHY and Factory-X funded by the Federal Ministry of Research, Technology and Space (BMFTR) and the Federal Ministry for Economic Affairs and Energy (BMWE).

References

[1] IDTA-Workstream "Specification of AAS". Specification of the Asset Administration Shell Part 1: Metamodel. Industrial Digital Twin Association e.V, Germany. doi:10.62628/IDTA.01001-3-1-1.

[2] Wagner, C., Grothoff, J., Epple, U., Drath, R., Malakuti, S., Gruner, S., Hoffmeister, M., Zimermann, P., 92017. The role of the Industry 4.0 asset administration shell and the digital twin during the life cycle of a plant, in: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol. 12.09.2017 - 15.09.2017. IEEE, pp. 1–8.

[3] Goetz, B., Schneider, M., Bauernhansl, T., 2024. Concept of Event-based AAS Transformation Engine. Procedia CIRP 130, 1244–1249.

[4] Filz, M.-A., Bosse, J.P., Herrmann, C., 2024. Digitalization platform for data-driven quality management in multi-stage manufacturing systems. J Intell Manuf 35 (6), 2699–2718.

[5] Quatrini, E., Costantino, F., Di Gravio, G., Patriarca, R., 2020. Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities. Journal of Manufacturing Systems 56, 117–132.

[6] Roselli, S.F., Dahl, M., Subramaniyan, M., Bekar, E.T., Skoogh, A., 2024. Integrating Machine and Quality Data for Predictive Maintenance in Manufacturing System, in: Thürer, M., Riedel, R., Cieminski, G. von, Romero, D. (Eds.), Advances in Production Management Systems. Production Management Systems for Volatile, Uncertain, Complex, and Ambiguous Environments, vol. 732. Springer Nature Switzerland, Cham, pp. 95–107.

[7] Himmelstoß, H., Risling, M., Rohde, J., Shi, D., Bauernhansl, T., 2025. Developing And Piloting A Procedure For Systematically Deriving The Digital Twin Based On The Asset Administration Shell. Hannover : publish-Ing.

[8] Risling, M., Oberle, M., Bauernhansl, T., 2024. Analyzing The Purpose And Technologies Of Digital Twins In Distributed Manufacturing: A Systematic Literature Review. Procedia Computer Science 232, 368–376.

[9] Russell, H., Himmelstoß, H., Böttinger, F., Kern, S., Bauernhansl, T., 2026. Causal Models and a Prototypical Implementation of a Pipeline for Causal Analysis in Manufacturing, in: Mas, F., Del Valle, C., Eynard, B., Rivest, L., Bouras, A. (Eds.), Product Lifecycle Management. PLM in the Age of Model-Based Engineering in Industry, vol. 772. Springer Nature Switzerland, Cham, pp. 182–190.

[10] Banerjee, A., Chen, C.-C., Hung, C.-C., Huang, X., Wang, Y., Chevesaran, R., 2020. Challenges and Experiences with MLOps for Performance Diagnostics in Hybrid-Cloud Enterprise Software Deployments, in: , Proceedings of the 2020 USENIX Conference on Operational Machine Learning (OpML '20). USENIX Association, 2020, [Berkeley, CA].

[11] Koukaras, P., 2025. Data Integration and Storage Strategies in Heterogeneous Analytical Systems: Architectures, Methods, and Interoperability Challenges. Information 16 (11), 932.

[12] Munappy, A.R., Bosch, J., Olsson, H.H., 2020. Data Pipeline Management in Practice: Challenges and Opportunities, in: Morisio, M., Torchiano, M., Jedlitschka, A. (Eds.), Product-Focused Software Process Improvement, vol. 12562. Springer International Publishing, Cham, pp. 168–184.

[13] Prince Canuma, 2022. MLOps: What It Is, Why It Matters, and How to Implement It. https://neptune.ai/blog/mlops.

[14] Coutinho, E.F., Carvalho Sousa, F.R. de, Rego, P.A.L., Gomes, D.G., Souza, J.N. de, 2015. Elasticity in cloud computing: a survey. Ann. Telecommun. 70 (7-8), 289–309.

[15] Herbst, N.R., Kounev, S., Reussner, R., 2020. Elasticity in Cloud Computing: What It Is, and What It IsNot, in: , 10th International Conference on Autonomic Computing (ICAC'13). Curran Associates, Inc., 2020, Red Hook, NY.

[16] Lehrig, S., Eikerling, H., Becker, S., 2015. Scalability, Elasticity, and Efficiency in Cloud Computing, in: Proceedings of the 11th International ACM SIGSOFT Conference on Quality of Software Architectures.

CompArch '15: Federated Events on Component-Based Software Engineering and Software Architecture, Montréal QC Canada. 04 05 2015 08 05 2015. ACM, New York, NY, USA, pp. 83–92.

[17] Mishra, S.K., Sahoo, B., Parida, P.P., 2020. Load balancing in cloud computing: A big picture. Journal of King Saud University - Computer and Information Sciences 32 (2), 149–158.

[18] Niu, Y., Liu, F., Li, Z., 2018. Load Balancing Across Microservices, in: IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, Honolulu, HI. 16.04.2018 - 19.04.2018. IEEE, pp. 198–206.

[19] Pahl, C., Jamshidi, P., Zimmermann, O., 2018. Architectural Principles for Cloud Software. ACM Trans. Internet Technol. 18 (2), 1–23.

[20] Wolff, E., 2018. Microservices: Grundlagen flexibler Softwarearchitekturen, 2., updated edition ed. dpunkt.verlag, Heidelberg, 1 p.

[21] Ainbinder, I., Temnikov, E., Allalouf, M., 2024. A Study Comparing Waiting Times in Global and Local Queuing Systems with Heterogeneous Workers. Applied Sciences 14 (9), 3799.

[22] Thallapally, N. Enhancing Distributed Systems with Message Queues: Architecture, Benefits, and Best Practices, in: , pp. 63–69.

[23] Bozkurt, M., Harman, M., Hassoun, Y., 2013. Testing and verification in service ‐ oriented architecture: a survey. Software Testing Verif & Rel 23 (4), 261 ‐ 313.

[24] Laskey, K.B., Laskey, K., 2009. Service oriented architecture. WIREs Computational Stats 1 (1), 101–105.

[25] Red Hat, Inc., 2020. What is service-oriented architecture (SOA)? https://www.redhat.com/en/topics/cloud-native-apps/what-is-service-oriented-architecture.

[26] Eclipse Foundation. eclipse-aaspe/server: C# based server for AASX packages. https://github.com/eclipse-aaspe/server.

[27] Eclipse Foundation. Eclipse BaSyx. https://eclipse.dev/basyx/.

[28] Fraunhofer IOSB. FAAAST Service. https://github.com/FraunhoferIOSB/FAAAST-Service.

[29] Stojanovic, L., Usländer, T., Volz, F., Weißenbacher, C., Müller, J., Jacoby, M., Bischoff, T., 2021. Methodology and Tools for Digital Twin Management—The FA3ST Approach. IoT 2 (4), 717–740.

[30] Soni, D., Makwana, A. (Eds.), 2017. A Survey on MQTT: A Protocol of Internet Of Things (IOT).

[31] Evans, B., Braun, S., Ulmer, J., Wollert, J., 2022. AAS Implementations - Current Problems and Solutions, in: 2022 20th International Conference on Mechatronics - Mechatronika (ME). 2022 20th International Conference on Mechatronics - Mechatronika (ME), Pilsen, Czech Republic. 07.12.2022 - 09.12.2022. IEEE, pp. 1–6.

[32] OASIS Standard, 2019. MQTT Version 5.0. https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html.

[33] Zertal, S., Harrison, P.G., Knottenbelt, W.J., 2001. Performance analysis of three implementation strategies for distributed lock management. IEE Proceedings - Computers and Digital Techniques 148 (4), 176–187.

[34] Industrial Digital Twin Association. Submodel Templates for AAS. https://github.com/admin-shell-io/submodel-templates. Accessed 1 March 2024.

Biography

Benjamin Goetz (*1986) works as a research associate in the research team “IT Architectures for Digital Manufacturing” at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He graduated in mechanical engineering at the Friedrich Alexander University Erlangen-Nuremberg with specialization in manufacturing technology.

Matthias Schneider (*1992) works as a research associate in the research team “IT Architectures for Digital Manufacturing” at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He graduated in mechatronics at the University Stuttgart with specializations in software engineering and industrial control systems.

Daniel Schel (*1986) works as a research team manager in the research area “IT-Systems” at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He graduated in software engineering at Heilbronn University of applied sciences and has managed and conducted a number of research projects as technical director and software architect, focusing on novel approaches to IT solutions in production.

Thomas Bauernhansl (*1969) has been Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA and the Institute of Industrial Manufacturing and Management IFF of the University of Stuttgart. His scientific interests are in the field of sustainable, digitalized, personalized, and biointelligent production.

8th Conference on Production Systems and Logistics

Self-Description Of Heterogeneous System Landscapes In Manufacturing Environments

Matthias Schneider[1], Benjamin Götz[1], Daniel Schel[1], Thomas Bauernhansl[1,2]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation, Stuttgart, Germany*
[2]*Institute of Industrial Manufacturing and Management, Stuttgart, Germany*

Abstract

Driven by digitalization, today's manufacturing environments are characterized by a growing number of heterogeneous devices, ranging from sensors and controllers to complex machines. Managing such heterogeneous system landscapes is getting increasingly complex and time-consuming and requires comprehensive knowledge of the characteristics and current state of each device. A key requirement for handling the complexity and heterogeneity of system landscapes is the availability of standardized, machine-readable self-descriptions of devices to support interoperability and lifecycle management. This work introduces a concept to combine multiple sources of information to address the challenges of creating self-descriptions in manufacturing environments with heterogeneous system landscapes. At its core, the concept leverages the Asset Administration Shell (AAS) as a standardized digital representation of devices, enabling structured integration of static manufacturer information and dynamic device-level information. A discovery process ensures the initial identification of devices, providing the foundation for subsequent enrichment with detailed information. The proposed Information Receiving Service handles multiple information sources, abstracts different access and authentication mechanisms, and aggregates access to multiple AAS servers. Complementing this, a Self-Description Service can run directly on devices and collect operational data from sources such as the operating system, monitoring tools, or container platforms. If manufacturer data is not natively available in AAS format, artificial intelligence (AI) can support the extraction and transformation of unstructured information, such as datasheets or manuals, into standardized submodels. The concept was prototypically implemented and integrated into the open-source project Eclipse Service Lifecycle Management.

Keywords

Self-Description; Heterogeneous System Landscapes; Digital Twin; Asset Administration Shell; Manufacturing Environments

1. Introduction

The digital transformation leads to an increased use of networked devices in manufacturing environments. Today's factories feature a diverse array of systems, encompassing a wide range of components, from sensors and controllers to complex machines that communicate and interact with one another. Managing the large amount and high degree of heterogeneity in the systems used – consisting of different vendors, software versions, and device types – continuously increases the effort required for management and maintenance [1] [2] [3]. Companies face the challenge of maintaining an overview of the entire system landscape. Information is scattered across different systems or manually maintained lists have become outdated [4]. The identification and cataloging of systems in a manufacturing environment, also known as discovery, is therefore an essential element. Once the existence of a system is generally known, a self-description of this

DOI: https://doi.org/10.15488/20972
ISSN: 2701-6277

system is required, so that the current state (e.g., configurations or firmware version) of the system can be assessed. A comprehensive self-description is the basis for deploying and maintaining software in such complex system landscapes, enabling updates and deployments to be supported by (partially) automated tools. Therefore, information about the systems, including their properties, functionalities, and current state, must be known and available in a standardized and machine-readable format.

In addition to operational issues, compliance with regulatory requirements is becoming an increasingly challenging task. Laws and regulations, such as the Cyber Resilience Act (CRA) [5], necessitate comprehensive knowledge of the system landscape and its current state. This knowledge provides transparency about the systems used, enabling companies to ensure that legal requirements are met and that security gaps can be identified early on and closed with updates. Upcoming regulatory initiatives such as the EU Digital Product Passport (DPP), which requires standardized, machine-readable lifecycle information for products and industrial equipment, further increase the need for interoperable self-descriptions of devices in manufacturing environments [6].

To address these challenges, this work examines how the integration of static manufacturer data and dynamic device-level information can contribute to the automated, interoperable self-description of heterogeneous system landscapes in manufacturing environments.

2. Related Work

2.1 Asset Administration Shell

Manufacturing environments contain a diverse range of systems from various manufacturers. A uniform and standardized description of device properties, as provided by the Asset Administration Shell (AAS), is therefore essential. The AAS forms the basis for the digital representation of assets and their capabilities in a standardized, structured, and machine-readable form. As a digital representation of a physical or virtual asset, the AAS provides structured information in so-called submodels, which contain various aspects of a device, e.g., nameplate information, technical data, capabilities, or state information. To ensure interoperability, the Industrial Digital Twin Association (IDTA) coordinates the development and harmonization of so-called submodel templates. These templates define semantically consistent structures and content for specific aspects of an asset, such as nameplate information or technical product properties. Standardization activities are carried out in close coordination with national and international standardization organizations such as DIN, IEC, and Platform Industry 4.0. [7] [8] [9]

In current research, several approaches already exist that focus on the automated creation and semantic enrichment of AAS. Xia et al. [10] demonstrate how large language models (LLMs) can be utilized to automatically generate AAS models from technical documents, such as data sheets. In another work [10], they employ a transfer learning approach, utilizing semantic fingerprints to efficiently convert proprietary data structures into standardized AAS submodels. Zhao et al. [11] present a semi-automated method that processes various data sources, such as CAD or PDF files, and generates initial AAS from them, which is useful in product-centric engineering phases but often insufficient to capture dynamic, device-level runtime information during operation phase. While these approaches help to bootstrap AAS creation, they typically address single data sources or isolated engineering views and do not provide an integrated concept for combining static manufacturer data with dynamic operational data from the devices. Model-driven methods, such as the model-based systems engineering (MBSE) approach proposed by Barbie et al. [12], bridge existing engineering models to AAS, yielding precise and scalable AAS instances when rich MBSE models exist. However, such methods rely on mature, centrally maintained models that are rarely available for legacy equipment and mixed-vendor environments.

Stock et al. [13] highlight the need for self-describing Cyber-Physical Production Systems (CPPS) as a prerequisite for higher levels of autonomy such as self-configuration and self-optimization, and explicitly

identify self-description as a core capability for the first level of autonomy. The role of the Asset Administration Shell as a standardized framework that facilitates interoperability and lifecycle management within these systems is emphasized. This work provides a strong motivational basis for the present contribution but does not address the concrete challenges of creating and maintaining self-descriptions of heterogeneous system landscapes in manufacturing environments.

The existing research shows that AAS-based digital twins and AI-driven AAS generation are technically feasible, but they typically focus on product-centric, design-time, or vendor-specific views. The remaining gap lies in a concept that integrates automated discovery, vendor information, and device-level operational information into standardized and comprehensive self-descriptions, while remaining applicable in heterogeneous system landscapes.

2.2 ID-Link

The Identification Link (ID-Link) is a globally unique reference mechanism that makes every device or device type uniquely addressable. The ID Link is defined in the international series of standards IEC 61406. Part 1 (IEC 61406-1:2022) describes the ID-link as a globally unique, machine-readable identifier that references digital information and can be made physically available on the device, e.g., in the form of a 2D code or an NFC tag. The syntax of an ID-link is URL-based, comprising the scheme, authority, path, fragment, and query components [14]. Part 2 (IEC 61406-2:2024) extends this specification with detailed rules for the syntactic structure and encoding of parameters, as well as the mapping of various information, such as device types, models, batches, or individual instances [15]. Unique identification is essential for the reliable management of heterogeneous system landscapes. It prevents devices from being mixed up or information from being assigned inconsistently. It provides the basis for consistently merging data from different sources and assigning it to a specific device instance or device type throughout its entire life cycle.

2.3 Discovery

Automatic discovery forms the basis for the initial and ongoing identification of systems in manufacturing environments. At the network layer, vendor-neutral protocols such as the Link Layer Discovery Protocol (LLDP) are widely used in industrial Ethernet networks [16] to detect directly connected devices, while in PROFINET environments the Discovery and Configuration Protocol (DCP) enables device discovery [17]. At the service level, mechanisms such as Multicast DNS (mDNS) together with DNS-SD support zero-configuration discovery of services without a central server [18], and OPC UA uses local and global discovery servers to register and locate endpoints [19]. In addition, protocols like SNMP provide discovery methods to identify and classify networked devices [20]. This brief overview illustrates that discovery in manufacturing environments can be realized through a combination of network- and service-level mechanisms, while numerous other, including proprietary and manufacturer-specific, discovery methods exist in practice. The challenge is to use the various discovery methods to discover as many devices as possible in heterogeneous system environments without exposing the complexity of these discovery methods to higher-level systems that utilize the information about the discovered devices.

2.4 Cyber Resilience Act

The CRA is an EU regulation that defines binding cybersecurity requirements for all products with digital elements offered in the EU, regardless of their scope of application. It was published in the Official Journal of the EU on 20 November 2024 and will come into full force with a transition period until the end of 2027 [21]. The aim is to establish a harmonized level of protection for digital products, thereby systematically reducing cybersecurity risks. Specifically, the CRA requires manufacturers to implement security by design and security by default. Manufacturers must, e.g., perform a comprehensive cybersecurity risk analysis and provide technical documentation, such as software bills of materials (SBOMs), as well as software updates. The regulation also provides for mandatory CE marking with cybersecurity certification, which can be carried out by the manufacturer's self-assessment or external testing bodies, depending on the risk class [21]

[22]. In the context of self-description, this demonstrates that a comprehensive, up-to-date, and standardized representation of devices is not only necessary for technical interoperability in heterogeneous production environments but also makes a significant contribution to fulfilling regulatory requirements.

3. Concept

The self-description of devices encompasses a wide range of information types that can be differentiated in terms of origin, structure, and relevance. Static information, mainly provided by the manufacturer, plays an essential role in this process. Such information can be either type-specific, e.g., in the form of general device models, technical specifications, or interface descriptions, or instance-specific, such as serial numbers or individual configuration values. Additionally, devices provide dynamic information that is continuously generated during operation or is only known by the device itself. Dynamic information includes, e.g., status information, process parameters, diagnostic messages, or performance indicators, which are often updated in real-time and can be used for higher-level monitoring, control, or optimization tasks. Only the combination of both types of information – static and dynamic – enables a consistent and comprehensive self-description of a device. For self-descriptions, a semantic and standardized description is required. Therefore, the AAS is utilized in the concept, as a standard for semantic description and the implementation of digital twins [23] [24]. Already existing submodel templates containing basic identification and nameplate information, technical properties, and operational and status-related data are essential for the self-description of devices. Depending on the application, more specific submodel templates may be required.

3.1 Architecture

Depending on the type of information, it is available from various sources, usually from the manufacturer or directly from the device. This information must be retrieved or made accessible from various sources and combined into a consistent self-description. Figure 1 illustrates the architecture concept. All devices must be managed using a Common Device Management (CDM). The term CDM summarizes software tools that manage devices in manufacturing environments. There are various solutions available from different vendors, such as Siemens Industrial Asset Hub [25], Hilscher netFIELD.io [26], OTbase from Langner [27], or the open-source project Eclipse Service Lifecycle Management [28]. The first step is to get knowledge about the existence of a device. A device can be either created manually in the CDM or discovered through drivers. Once a device has been identified, an AAS is created for it. The asset ID of the AAS relates to the device ID (e.g., ID Link). The device AAS forms the basis for a device's self-description and requires information from various sources to be filled in. It contains all submodels of a device, each describing specific aspects of the device. If a manufacturer provides information about a device via AAS, this information can be retrieved directly from the manufacturer's AAS services using the Information Receiving Service (IRS). The basis for retrieving this information is the ID Link, which ensures a unique assignment between a specific device and the associated manufacturer information. In this case, the ID Link corresponds to the asset ID of the AAS. If this relation between ID Link and asset ID is not available, query methods as described in the Specification of the Asset Administration Shell, Part 2: Application Programming Interfaces [29], can be used to search for relevant information, e.g., by using the product name or serial numbers. If information is not provided by the manufacturer in AAS format, the creation of submodels can be supported with the help of the AI Submodel Assistant. This tool extracts information from sources such as manuals and product websites, matches it to existing submodel templates and saves the user manual work. The user only needs to review the generated submodels and correct or expand them as necessary, without needing to start the submodel creation from scratch. Further information, especially dynamic information, is provided directly by the device via the Self-Description Service (SDS). The SDS can be installed automatically on the device via CDM or manually during device onboarding. The submodels of the SDS are linked to the

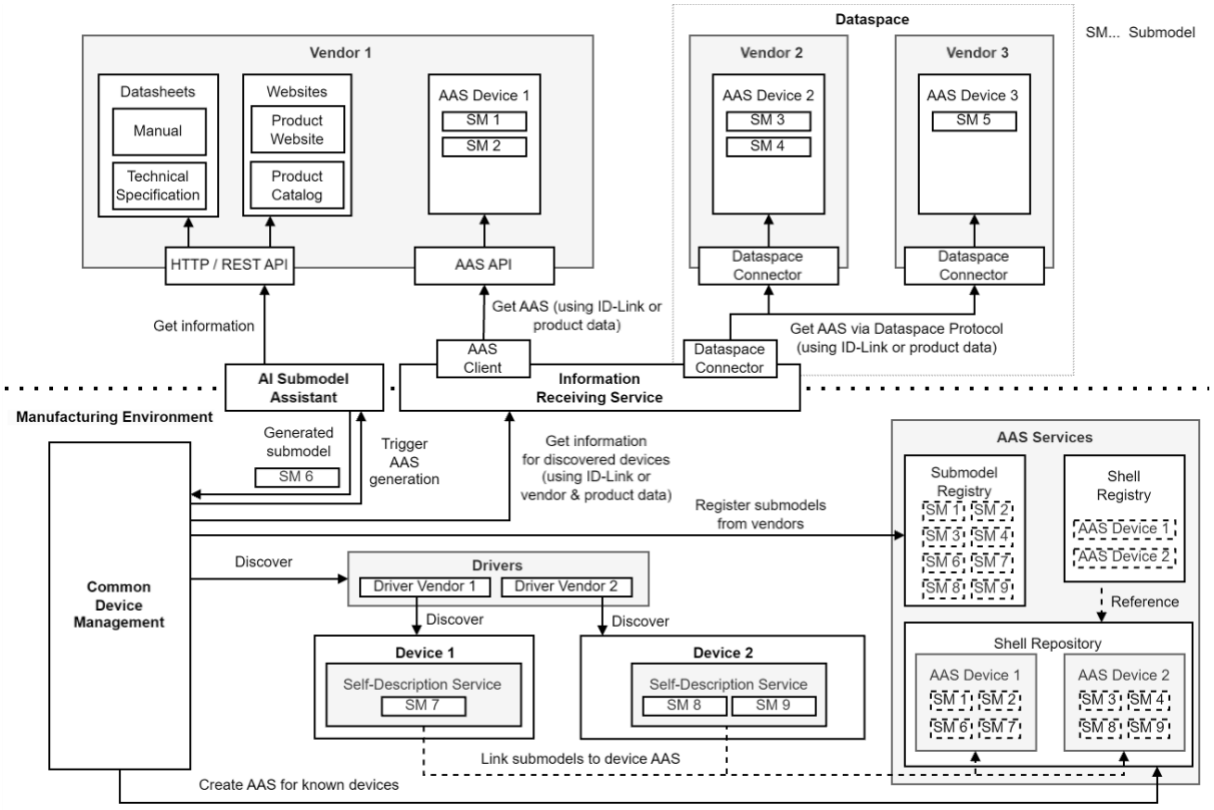


Figure 1: Architecture

corresponding device AAS. The following sections provide a detailed description of the concept's main components. As a result of combining these components, a comprehensive AAS is generated for each device, combining the relevant submodels from different information sources. This ensures that the device's self-description is not only comprehensive and consistent but also interoperable and compliant with standards.

3.2 Discovery

In addition to describing the devices, the first essential question to be answered is: Which devices exist in a manufacturing environment? As shown in Figure 2, device discovery is performed using a discovery process that automatically identifies devices over the network. The process is based on manufacturer-specific or protocol-specific drivers that can discover devices of a specific type or manufacturer, or via a specific protocol. As part of the discovery process, a device usually returns only a small amount of essential information, such as the manufacturer, product name, or serial number. Some newer devices also provide an identification link (ID Link). This basic information serves as the starting point for retrieving more detailed information from manufacturers. The uniform interface of the discovery drivers enables integration into higher-level systems, such as a CDM. This ensures that device discovery is consistent, regardless of the manufacturer or protocol, and can be seamlessly integrated into existing management systems.

3.3 Information Receiving Service

An AAS server of a manufacturer can be directly accessible via the internet or accessible via federated data infrastructures such as dataspaces. Access can be either public or protected, requiring authorization. In the case of dataspaces, access is done via so-called connectors and the Dataspace Protocol (DSP). Dataspace technology provides manufacturers with additional options for data governance and simplifies authentication and authorization from the perspective of the data consumer. Data consumers have an identity in the dataspace, which allows the respective manufacturers to grant them access rights to their data. There are currently no fully established dataspaces in the manufacturing area. Dataspaces are still being developed and driven by the Industrial Dataspace Association (IDSA) [30] and projects such as Factory-X [31] and Catena-X [32]. Therefore, in the short and medium term, data consumers will need to manage multiple access options (e.g., access via pure AAS methods or access to AAS in combination with DSP). The Information Receiving

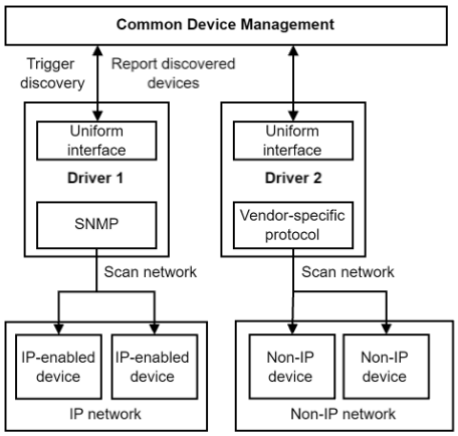


Figure 2: Discovery concept

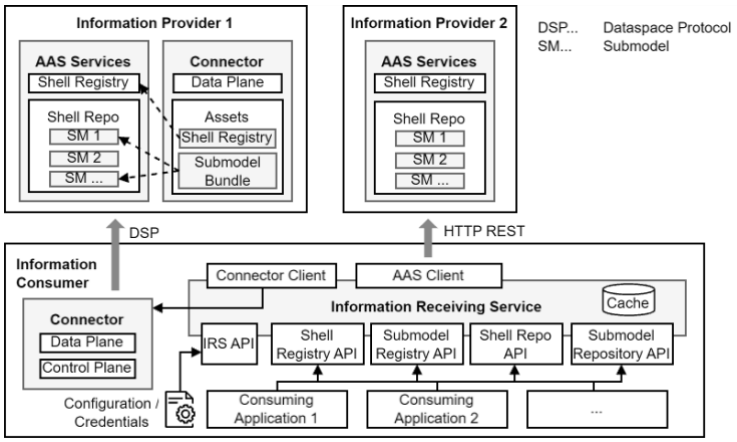


Figure 3: Information Receiving Service

Service (IRS) depicted in Figure 3 was designed to simplify data consumption in this context. It abstracts the heterogeneity of data sources by providing the same interfaces as an AAS service but can internally aggregate access to multiple AAS services or dataspace connectors. This means that requests to the IRS are sufficient to access information from different manufacturers. The IRS aggregates the various AAS sources and abstracts the transport protocol for transferring AAS. Credentials only need to be managed in one place and not by multiple data consumers. This reduces the integration effort for data consumers. Furthermore, AAS information can be cached by the IRS if the same information needs to be retrieved frequently or if the manufacturer's AAS services are temporarily unavailable.

3.4 AI Submodel Assistant

In the best-case scenario, all information is provided directly by the device or manufacturer in AAS format. Unfortunately, in practice, this is often not yet the case, as many manufacturers publish information in conventional form – such as data sheets, manuals, or websites – and are gradually offering information via AAS. In such cases, it is possible to manually add additional submodels to the device's self-description. Creating submodels is time-consuming and requires in-depth knowledge of the AAS standard. An AI Submodel Assistant, as shown in Figure 4, can be used to support this process. It can analyze existing information sources and automatically fill standardized submodel templates with the extracted data using a large language model (LLM). Depending on the use case and its requirements, this process can be fully automated or interactive, with a user reviewing, supplementing, and approving the proposed submodels. The AI assistance simplifies the integration of devices without native AAS support and expands the self-description.

3.5 Self-Description Service

The information from the manufacturer is supplemented by data collected directly from the device through a Self-Description Service. The concept of the service is shown in Figure 5. It runs directly on the devices, has a modular structure, and can make various data sources available directly on the device as submodels. Possible data sources include operating system information, configuration files, monitoring exporters such as Node Exporter [33], or container metadata from container engines like Docker. This creates a flexible architecture that integrates both static information provided by the manufacturer and dynamic, device-related data.

The Self-Description Service uses the AAS REST API as a standardized interface and data format. In particular, the service provides an API compliant with the IDTA's Submodel Repository Open API specification [34]. The submodels are generated dynamically on request. The generation is based on parameterized templates that contain references to data sources, which are replaced with actual values from

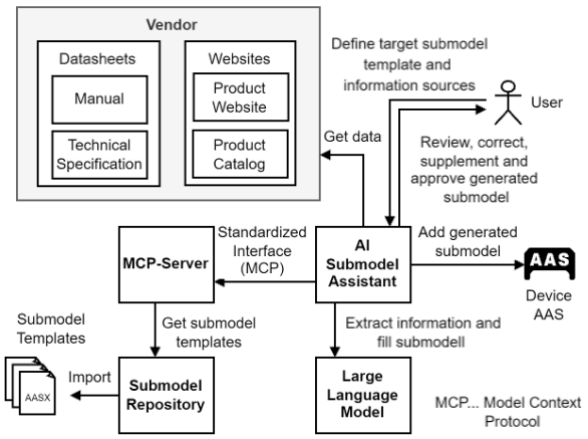


Figure 4: AI Submodel Assistant

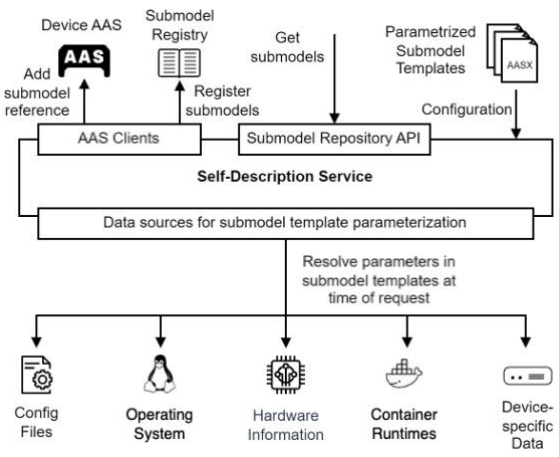


Figure 5: Self-Description Service

the data source when the submodel is requested. In contrast to IDTA submodel templates, which define the structural design of a submodel, the templates in this context are more like parameterized submodel templates. They contain variables and references to data sources that are dynamically resolved at runtime to generate fully populated submodels. Integrated clients for AAS Registry, AAS Repository, and Submodel Registry enable the submodels provided by the Self-Description Service via its Submodel Repository API to be automatically registered and referenced in an existing AAS of a device. Such an existing AAS may already contain static information retrieved from the manufacturer. The Self-Description Service supplements the AAS with additional information from the device, particularly dynamic information.

To make configuration as simple as possible and enable less technically experienced users to easily create parameterized submodel templates, the definition is done via AASX files. The templates can be created using graphical tools such as the Eclipse AASX Package Explorer [35]. The service features a modular structure that supports various data sources for parameterization, including configuration files (e.g., YAML, JSON), the extraction of operating system information from shell commands, monitoring tools such as Prometheus Exporter, and container environments like Docker. Through the modular software design, further data sources can be easily integrated. By utilizing the template mechanism, different data sources can be combined flexibly and efficiently. The dynamic generation of complex submodels with many parameterizations can be computationally time-consuming. To prevent delays or excessive load on the device caused by the Self-Description Service, caching can be configured for the data sources. Depending on the configuration, the values of the data sources are then stored in the cache and updated only when necessary (e.g., if the data becomes outdated). Since the device can also have sensitive information (e.g., passwords or critical configuration parameters), parameterized submodel templates can be marked as confidential, allowing them to be accessed only by authorized users. Additionally, semantic IDs can be configured that should not be output in the submodel templates, preventing accidental misconfiguration. In addition to explicitly referencing data sources in the parameterized submodel templates (e.g., retrieve property XY from file Z), data sources can also register semantic IDs for which they can provide values (e.g., current CPU utilization). When resolving the parameterized submodel templates, the semantic IDs can be checked to see if data sources are registered that can provide values for these semantic IDs. In this case, the data source can be requested to provide the value.

4. Implementation and Evaluation

The Self-Description Service [36] and the Information Receiving Service [37] were implemented as standalone components, can be used independently in other contexts, and are available as open source. The AAS-specific parts were implemented using the Java Server SDK from Eclipse Basyx [38]. Drivers for discovery were developed using an open source SDK initially published by Siemens [39]. Some of the

drivers were made available by manufacturers as part of the Factory-X project. The AI-assisted generation of submodels is part of the concept, but the development of the AI Submodel Assistant component was not the focus of this work. It is based on existing preliminary work [40] [41] and was tested with a prototypical implementation of the Fraunhofer IPA. The publication and open-source release of this development are in preparation. For the overall integration and evaluation of the concept, the individual components were integrated into the open-source project Eclipse Service Lifecycle Management, which acts as a CDM in this context [28].

An evaluation was performed using a demonstrator comprising various device types from different manufacturers (e.g., Siemens PLC, Festo Servo Drive Controller, Hilscher IO-Link Wireless Master, OnLogic Industrial NUC, and industrial-grade Raspberry Pis). These devices represent a heterogeneous system landscape in a manufacturing environment. Discovery was tested with devices from Festo, Hilscher, Sick, and Siemens, and the retrieval of static manufacturer-specific information was tested with the AAS development servers of these manufacturers. For dynamic information, the Self-Description Service was deployed on devices with Linux-based operating systems. For the various device types, it was thus possible to demonstrate that, after an initial discovery of the devices, a comprehensive self-description could be created based on AAS, with static information retrieved from the manufacturer and dynamic information provided directly from the device.

5. Conclusion

This work presents a novel concept for creating comprehensive, machine-readable, and standardized self-descriptions of heterogeneous system landscapes in manufacturing environments by integrating different information sources, leveraging the AAS as a standardized digital representation for the devices. The key scientific contribution lies in the combination of static manufacturer data with dynamic, device-level data through modular, interoperable services – specifically the Information Receiving Service and the Self-Description Service. The prototypical implementation within the Eclipse Service Lifecycle Management demonstrates the practical feasibility of the approach across diverse industrial devices and validates the automated discovery and aggregation mechanisms.

Despite these advances, several challenges remain open. Manufacturer information is often not yet available in AAS format. Manually creating submodels for such missing information is very time-consuming. AI can currently support the creation of submodels, but it cannot yet fully automate this process. Beside generating the submodels, further standardization of submodel templates for the Asset Administration Shell remains an ongoing task. To comprehensively represent the diverse aspects of devices for various use cases, submodel templates must be continuously developed and reach a broad consensus among users and organizations. Legacy devices often do not have a unique identifier, which leads to a certain degree of inaccuracy when searching for manufacturer information. For newer and future devices, this will no longer be an issue, as regulations such as the Digital Product Passport sets requirements regarding this and standards like ID-Link already address this problem.

Future research will build on the presented self-description concept to enable dynamic and automated software deployment in manufacturing environments. For this, it is essential to know the properties of devices within the system landscape to align them with software requirements. Additionally, up-to-date device status information is essential for adapting deployments in response to changes in the system landscape.

Acknowledgements

This research was conducted as part of the Factory-X research project, funded by the Federal Ministry for Economic Affairs and Energy (BMWE) of Germany.

References

[1] Khan, W.Z., Rehman, M.H., Zangoti, H.M., Afzal, M.K., Armi, N., Salah, K., 2020. Industrial internet of things: Recent advances, enabling technologies and open challenges. Computers & Electrical Engineering 81.

[2] Arjun, K.P., Smitha, V., 2025. Smart Factories in the Era of Industry 4.0: Convergence of IoT and Cyber-Physical Systems for Intelligent Automation, Connectivity, and Resilient Manufacturing, in: Sonwaney, V., Gupta, S.K., Nayal, K., Nirmal, D. (Eds.), Proceedings of the International Conference on Operations & Supply Chain Management 2025 (ICOSCM 2025), vol. 364. Atlantis Press International BV, Dordrecht, pp. 261–270.

[3] Shi, Z., Xie, Y., Xue, W., Chen, Y., Fu, L., Xu, X., 2020. Smart factory in Industry 4.0. Syst Res Behav Sci 37 (4), 607–617.

[4] Pakkala, D., Kääriäinen, J., Mätäsniemi, T., 2024. Improving efficiency and quality of operational industrial production assets information management in customer–vendor interaction. Journal of Industrial Information Integration 41, 100644.

[5] 2024. Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act): CRA, in: , Official Journal of the EU.

[6] European Union, 2024. EU's Digital Product Passport: Advancing transparency and sustainability. https://data.europa.eu/en/news-events/news/eus-digital-product-passport-advancing-transparency-and-sustainability. Accessed 20 March 2026.

[7] IDTA-Workstream "Specification of AAS". Specification of the Asset Administration Shell Part 1: Metamodel. Industrial Digital Twin Association e.V, Germany. doi:10.62628/IDTA.01001-3-1-1.

[8] Deutsche Kommission Elektrotechnik, Elektronik, Informationstechnik, 2024. DIN EN IEC 63278-1 (VDE 0810-781), Asset administration shell for industrial applications. Part 1, Asset administration shell structure (IEC 63278-1:2023), Deutsche Fassung EN IEC 63278-1:2024 ed. VDE VERLAG GMBH; DIN Media GmbH, Berlin, 79 Seiten.

[9] Industrial Digital Twin Association. Submodel Templates for AAS. https://github.com/admin-shell-io/submodel-templates. Accessed 1 March 2024.

[10] Xia, Y., Xiao, Z., Jazdi, N., Weyrich, M., 2024. Generation of Asset Administration Shell with Large Language Model Agents: Toward Semantic Interoperability in Digital Twins in the Context of Industry 4.0.

[11] Zhao, J., Vogel-Heuser, B., Bi, F., Höfgen, J., Ocker, F., Vojanec, B., Markert, T., Kraft, A., 2023. A Semi-Automatic Approach for Asset Administration Shell Creation from Heterogeneous Data. IFAC-PapersOnLine 56 (2), 3673–3679.

[12] Barbie, P., Pollom, A., Fischer, R.-P., Becker, M., 2025. Automated Generation of Standardised Digital Twins Based on MBSE Models, in: Proceedings of the 13th International Conference on Model-Based Software and Systems Engineering. 13th International Conference on Model-Based Software and Systems Engineering, Porto, Portugal. 26.02.2025 - 28.02.2025. SCITEPRESS - Science and Technology Publications, pp. 75–84.

[13] Stock, D., Bauernhansl, T., Weyrich, M., Feurer, M., Wutzke, R., 2020. System Architectures for Cyber-Physical Production Systems enabling Self-X and Autonomy, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria. 08.09.2020 - 11.09.2020. IEEE, pp. 148–155.

[14] Deutsche Kommission Elektrotechnik, Elektronik, Informationstechnik, 2023. DIN EN IEC 61406-1 (VDE 0810-406-1), Identification link. Part 1, General requirements (IEC 61406-1:2022), Deutsche Fassung EN IEC 61406-1:2022 ed. VDE VERLAG GMBH; Beuth Verlag GmbH, Berlin, 53 Seiten.

[15] Deutsche Kommission Elektrotechnik, Elektronik, Informationstechnik, 2024. DIN EN IEC 61406-2 (VDE 0810-407), Identification link. Part 2, Types/models, lots/batches, items and characteristics (IEC 61406-2:2024), Deutsche Fassung EN IEC 61406-2:2024 ed. VDE VERLAG GMBH; DIN Media GmbH, Berlin, 39 Seiten.
[16] IEEE Standard for Local and metropolitan area networks - Station and Media Access Control Connectivity Discovery. IEEE, Piscataway, NJ, USA. doi:10.1109/IEEESTD.2016.7433915.
[17] PROFINET University. DCP – Discovery and Configuration Protocol. https://profinetuniversity.com/naming-addressing/profinet-dcp. Accessed 23 October 2025.
[18] Internet Engineering Task Force (IETF), 2013. Multicast DNS. https://datatracker.ietf.org/doc/html/rfc6762. Accessed 23 October 2025.
[19] OPC Foundation. Discovery Service Set. https://reference.opcfoundation.org/Core/Part4/v104/docs/5.4. Accessed 23 October 2025.
[20] Internet Engineering Task Force (IETF), 1999. Simple Service Discovery Protocol/1.0. https://datatracker.ietf.org/doc/html/draft-cai-ssdp-v1-03. Accessed 23 October 2025.
[21] 2024. Verordnung (EU) 2024/2847 des Europäischen Parlaments und des Rates vom 23. Oktober 2024 über horizontale Cybersicherheitsanforderungen für Produkte mit digitalen Elementen und zur Änderung der Verordnungen (EU) Nr. 168/2013 und (EU) 2019/1020 und der Richtlinie (EU) 2020/1828 (Cyberresilienz-Verordnung): CRA, in: , Official Journal of the EU.
[22] Federal Office for Information Security. Cyber Resilience Act: Cybersecurity in the EU. https://www.bsi.bund.de/dok/cra-en. Accessed 15 October 2025.
[23] Beden, S., Cao, Q., Beckmann, A., 2021. Semantic Asset Administration Shells in Industry 4.0: A Survey, in: 2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS). 2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS), Victoria, BC, Canada. 10.05.2021 - 12.05.2021. IEEE, pp. 31–38.
[24] Wagner, C., Grothoff, J., Epple, U., Drath, R., Malakuti, S., Gruner, S., Hoffmeister, M., Zimermann, P., 92017. The role of the Industry 4.0 asset administration shell and the digital twin during the life cycle of a plant, in: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol. 12.09.2017 - 15.09.2017. IEEE, pp. 1–8.
[25] Siemens AG. Industrial Asset Hub: INDUSTRIAL DEVICE MANAGEMENT. https://www.siemens.com/en-us/products/industrial-asset-hub/. Accessed 20 March 2026.
[26] Hilscher Gesellschaft für Systemautomation mbH. Device Management made easy. https://www.hilscher.com/lp/device-management-made-easy. Accessed 20 March 2026.
[27] Langner, Inc. OTbase: OT asset inventory done right. https://otbase.com. Accessed 20 March 2026.
[28] Eclipse Foundation AISBL. Eclipse Service Lifecycle Management. https://projects.eclipse.org/projects/dt.slm. Accessed 29 February 2024.
[29] IDTA-Workstream "Specification of AAS". Specification of the Asset Administration Shell Part 2: Application Programming Interfaces. Industrial Digital Twin Association e.V, Germany. doi:10.62628/IDTA.01002-3-1-1.
[30] International Data Spaces Association. https://internationaldataspaces.org. Accessed 15 October 2025.
[31] Factory-X: The digital ecosystem. https://factory-x.org. Accessed 15 October 2025.
[32] Catena-X: Your Automotive Network. https://catena-x.net. Accessed 15 October 2025.
[33] Node exporter. https://github.com/prometheus/node_exporter. Accessed 15 October 2025.
[34] Industrial Digital Twin Association (IDTA), 2025. DotAAS Part 2 | HTTP/REST | Submodel Repository Service Specification. https://app.swaggerhub.com/apis/Plattform_i40/SubmodelRepositoryServiceSpecification/V3.1.1_SSP-003. Accessed 15 October 2025.

[35] Eclipse AASX Package Explorer and Server. https://projects.eclipse.org/projects/dt.aaspe. Accessed 23 October 2025.
[36] Resource Self-Description Service. https://github.com/eclipse-slm/resource-self-description-service. Accessed 15 October 2025.
[37] Information Receiving Service (IRS). https://github.com/eclipse-slm/information-receiving-service. Accessed 15 October 2025.
[38] Eclipse Foundation. Eclipse BaSyx. https://eclipse.dev/basyx/.
[39] Asset Link SDK. https://github.com/industrial-asset-hub/asset-link-sdk. Accessed 15 October 2025.
[40] Shi, D., Liedl, P., Bauernhansl, T., 2024. Interoperable information modelling leveraging asset administration shell and large language model for quality control toward zero defect manufacturing. Journal of Manufacturing Systems 77, 678–696.
[41] Shi, D., Meyer, O., Oberle, M., Bauernhansl, T., 2025. Dual data mapping with fine-tuned large language models and asset administration shells toward interoperable knowledge representation. Robotics and Computer-Integrated Manufacturing 91, 102837.

Biography

Matthias Schneider (*1992) works as a research associate in the research team "IT Architectures for Digital Manufacturing" at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He graduated in mechatronics at the Universität Stuttgart with specializations in software engineering and industrial control systems.

Benjamin Goetz (*1986) works as a research associate in the research team "IT Architectures for Digital Manufacturing" at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He graduated in mechanical engineering at the Friedrich Alexander Universität Erlangen-Nürnberg, specializing in manufacturing technology.

Daniel Schel (*1986) works as a research team manager in the research area "IT-Systems" at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He graduated in software engineering at Heilbronn University of Applied Sciences and has managed and conducted several research projects as technical director and software architect, focusing on novel approaches to IT solutions in production.

Thomas Bauernhansl (*1969) is the Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA and the Institute of Industrial Manufacturing and Management IFF of the University of Stuttgart. His scientific interests are in the field of sustainable, digitalized, personalized, and biointelligent production.

8th Conference on Production Systems and Logistics

Future Competitiveness In Industrial Wood Processing: A Survey-Based Analysis Of Economic, Technological and Ecological Developments

D. Roppelt[1], J. Schmidt[1], B. Thorenz[1], .F. Döpper[1, 2]
[1]*University of Bayreuth, Chair of Manufacturing and Remanufacturing Technology, Bayreuth, Germany*
[2]*Fraunhofer IPA Project Group Process Innovation, Bayreuth, Germany*

Abstract

The German sawmill industry represents a key component of the national wood-processing sector and provides essential raw materials for a wide range of economic activities. At the same time, the industry is increasingly affected by climatic changes, variations in wood quality, growing demand, and continuous technological and economic developments. Against the backdrop of forest transformation and the transition toward sustainable resource use, the analysis of the sawmill industry's competitiveness and adaptability has gained importance. In addition, there is growing interest in evaluating market trends in tool development to promote technological innovation under changing forest and wood structure conditions. The findings offer valuable insights for policymakers, industry stakeholders, and researchers in developing long-term strategies for a resilient and sustainable sawmill industry. To date, comprehensive empirical studies systematically differentiating technological, economic, and ecological factors by company size have been lacking. This study aims to identify differences between small, medium-sized, and large sawmills regarding current challenges, innovation needs, and adaptation strategies. The data are based on a nationwide quantitative survey conducted 31 German sawmills of different sizes (response rate: 11,4%), assessing market conditions, technological developments, and ecological influences. Results indicate that energy efficiency (72%), extended tool life (58%), and improved machine stability (45%) are considered highly relevant. Furthermore, ongoing forest conversion significantly impacts the raw material base, as 90% of respondents report an increasing share of damaged wood, while 19% identifying new market opportunities through the use of alternative tree species. The study demonstrates that the long-term competitiveness of the German sawmill industry depends on the integration of technological innovation, ecological adaptability, and supportive economic policy frameworks. Strengthening small and medium-sized sawmills is particularly important to ensure regional wood processing, maintain employment in rural areas, and achieve ecological benefits through reduced transport distances.

Keywords
Market Study; Sawmill Industry; Machining of Wood; Circular Saw Blade; Resource and Energy Efficiency

1. Introduction

The timber industry is currently undergoing profound change, driven by environmental, economic and social factors. Climate change in particular has had a significant impact on forests and timber production in recent years. Increasing drought, pest infestation and extreme weather events are leading to rising forest damage and changing the dynamics of timber use. Whereas storms used to be the main cause of damaged timber, insect infestation is now a key factor. [1]

DOI: https://doi.org/10.15488/20973
ISSN: 2701-6277

Active forest conversion towards more mixed and deciduous forests has also become significantly more urgent in recent years. Spruce in particular, which suffered considerable losses during the last inventory period, is increasingly being replaced by more climate-stable tree species. These changes in tree species composition also have an impact on wood processing. Sawmills must adapt to changes in the range of raw materials and adjust their production to the increasing diversity and proportion of coniferous wood. [2]

After years of rising demand for wood, particularly from the construction industry and for energy use, a significant decline in timber harvesting has been observed since 2024. The reasons for this are lower quantities of damaged wood, a certain degree of market saturation and weather-related restrictions. Overall, timber harvesting was significantly lower, especially for spruce, oak and beech, while pine was the only wood species to record a slight increase. [3,4]

At the same time, a significant increase in timber prices has been observed since mid-2025. The producer price index for timber products from state forests stood at 115.4 points in August 2025 (2015 = 100), which is 1.7% higher than in the previous month and 10.1% above the previous year's level. The price increase for logs and log sections was equally pronounced (from 102.2 in August 2024 to 114.6 in August 2025). [5]

The sawmill industry, as the link between forestry and the wood processing industry, is therefore facing a complex situation: declining timber volumes are coinciding with rising raw material prices, growing sustainability requirements and stricter legal framework conditions. Added to this are the external repercussions of geopolitical events and developments, as well as internal challenges such as a shortage of skilled workers, rising production costs and increasing bureaucracy. [6,7]

These developments illustrate that the sawmill industry is caught between ecological responsibility, economic pressure and technological change. Against this backdrop, this paper analyses current economic, technological, and ecological developments in the German sawmill industry, with particular attention to company-size-specific differences and derives implications for future innovation in sawing tool technologies under changing raw material conditions. The results should contribute to identifying sustainable, resource-efficient solutions for industrial wood processing that are adapted to changing raw material conditions. The paper is structured as follows: Section 2 provides an overview of the state of the art, including the German sawmill industry, sawing technologies, and fundamentals of market research. Section 3 outlines the approach and study design, while Section 4 presents the results and discussion. Section 5 concludes with key findings and outlook.

2. State of the Art

2.1 The German Sawmill Industry: National Structure and International Perspectives

The German sawmill industry forms the central link between forestry and wood processing. It handles the primary processing stage of round timber harvested in the forest. According to data from the Federal Statistical Office, around 31.1 million solid cubic meters (m³, equivalent to the German unit Festmeter, Fm) of raw timber were processed in 2024. This figure is based on data reported by sawmills subject to reporting requirements (annual production of 5,000 m³ or more; since 2009, companies with 10 or more employees). [8,9] By way of comparison, the raw material monitoring system for timber calculated an estimated felling volume of 39.147 million solid cubic metres for 2018. [10,11]

The industry encompasses a wide range of operating sizes, from small village sawmills that primarily serve regional demand to large industrial plants with an international market focus. According to the 2018 assessment of sawn timber output and production volumes conducted as part of the Wood Raw Materials Monitoring [10], the total population comprised 1,761 active sawmills, representing a decline of 14.9% compared to 2015. The 500–999 m³ (−34.4%) and 1,000–2,499 m³ (−41.6%) cutting size classes were particularly hard hit, while total roundwood cutting increased by 8.7% compared to 2015. [10,11]

Figure 1 shows the distribution of sawmills across cutting-capacity classes in Germany. Notably, small sawmills (<5,000 m³) account for 72% of all enterprises but produce only 3% of the total sawn timber volume. In contrast, the largest companies, representing approximately 6% of all sawmills, are responsible for 83% of total sawn timber production.

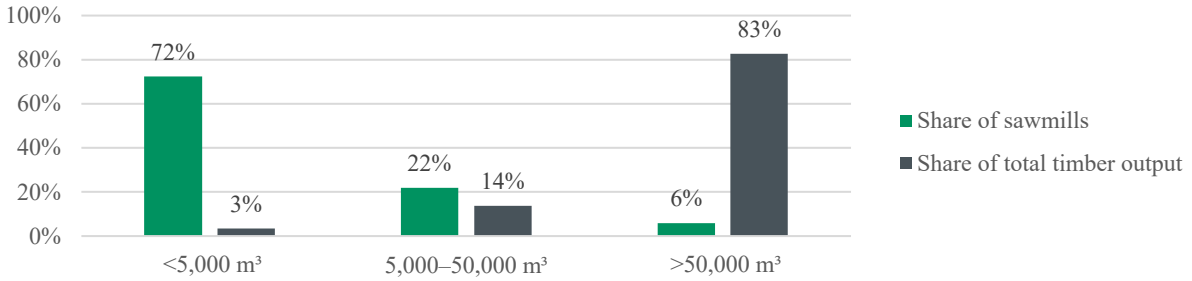


Figure 1: Distribution of German sawmills by annual timber cutting volume [10,11]

A total of 19,908 people were employed in the 332 sawmills, planing mills and wood impregnation plants with more than 10 employees. In 2024, these enterprises generated a turnover of around 7.47 billion euros [12]. The main products of sawmills are sawn timber for construction, windows, stairs and packaging. By-products such as sawdust and residual wood are either further processed industrially or used for energy production [13]. Overall, the German sawmill industry demonstrates a high degree of adaptability, but faces the ongoing challenge of balancing production, sustainability and market requirements while ensuring a sufficient supply of skilled workers [7]. A view of the European sawmill industry shows that it is saturated, highly production-oriented, and characterized by intense price competition. Material and labor costs continue to constrain competitiveness, while value-added products are becoming increasingly important. [14,15] In this highly competitive European environment, digital transformation is becoming increasingly important for small and medium-sized sawmills as a key driver of more efficient production processes and strategic development [16]. Globally, wood product supply chains are shaped by market integration, certification, traceability, and logistics. Europe and North America dominate production and consumption, while China has rapidly expanded as a major processor and exporter of value-added wood products. Despite advances in digital tools and governance frameworks, supply chains remain fragmented, smallholders are often excluded, and digital traceability is still limited in developing regions [17]. The international perspective highlights the relevance of the study beyond Germany. Moreover, the review of previous studies shows that there is little differentiated insight into the challenges faced by sawmills of different sizes, making this work particularly valuable for identifying sustainable solutions in industrial wood processing.

2.2 Sawing Technologies

In wood processing, sawing refers to the material-removing separation of a log with minimal energy consumption and cutting loss. Sawing methods can be classified into three main categories according to the type of motion: rotating saws (e.g., circular saws), continuous saws (e.g., band or chain saws), and reciprocating saws (e.g., gang or frame saws). [18] Circular saws are the most versatile method and are used for both longitudinal and crosscuts as well as precision cuts in wood materials. The rotating saw blade is simultaneously fed over the workpiece, allowing the cutting edges to completely sever the material. Blades are usually made of low-alloy steel, with cutting edges made of carbide, stellite, or diamond-based cutting materials, such as polycrystalline diamond, abbreviated as PCD, or diamond coatings produced by chemical vapour deposition (CVD) Diameters range from 200 mm to over 2.5 m, and blade thickness varies between 0.8 and 6.0 mm. Different tooth forms, such as mouse-tooth, pointed, or wolf-tooth, allow for varying cutting qualities. [18] Band saws can be divided into trim band saws for workshop and carpentry use, and block band saws as the main cutting units in sawmills. Block band saws are often integrated into transfer lines, can produce multiple parallel cuts, and allow flexible processing through special feed or return mechanisms. The thin saw bands are made of low-alloy steel, welded or brazed at the ends, and fitted with swaged or stellited teeth to increase durability and cutting quality. [18,19] Gang or frame saws are primarily used in sawmills

for producing multiple parallel cuts in a single log. They combine an oscillating cutting motion of the blades with a continuous or intermittent feed. The blades are made of carbon or low-alloy steel, can be stellited or hard-chrome plated, and are mounted in the frame to lift during the return stroke. [18,19] Together, these methods form the backbone of modern sawmilling, with each technique offering specific advantages in terms of cut quality, wood species, and production volume.

2.3 Fundamentals of Market Research

Market research is a systematic process of collecting, analyzing, and interpreting market-related data to support marketing decisions. It aims to identify opportunities and risks, monitor marketing activities, and reduce uncertainties. The process is commonly structured into five phases, the "5 D's of Market Research" (see Figure 2). In the definition phase, the research problem, objectives, and target population are established. The design phase involves selecting research methods, variables, and approaches, which can be exploratory, descriptive, or causal. [20]

Figure 2: Five-phase model of the market study [20]

Data collection may rely on primary research (original data) or secondary research (existing data) [21]. Primary research can be qualitative (e.g., interviews) or quantitative (standardised surveys), conducted as cross-sectional or longitudinal studies, and may also include experiments [20]. After data collection, the information is checked, coded, analyzed, and interpreted. The findings are documented and presented in a manner that ensures usability for strategic decisions. [21,20]

3. Approach and Study Design

For the primary analysis, a standardized online survey was conducted using Tivian Enterprise Feedback Suite (EFS). The questionnaire included open-ended and closed questions (single choice, multiple choice, matrix) and primarily collected quantitative data. General company information enabled classification by company size, allowing for detailed analyses in some cases. The survey link was accessible for four weeks. Thematically structured questions addressed economic, technological, and ecological aspects of the industry.

The sampling frame was derived from publicly available industry directories. From this frame, 272 sawmills were randomly selected using a stratified random approach to ensure balanced representation across company sizes and regions, covering the full spectrum from small village sawmills to large international enterprises. Due to practical constraints, including incomplete contact information and limited time resources, a survey of all ~1,700 German sawmills proved infeasible; 100 randomly selected companies were additionally contacted by phone to boost the response rate. Comparison with national statistics confirms that the participating sawmills reflect the structure of the German sawmill industry in terms of company size and production capacity, supporting the representativeness of the sample. The questionnaire was developed based on current industry challenges and reviewed by a saw blade manufacturer to ensure practical relevance. Potential biases, including question wording, order effects, and socially desirable responses, were addressed, and the questionnaire was pretested with three companies to ensure clarity, neutrality, and comprehensibility.

Of the 272 contacted companies, 31 completed the survey, corresponding to a response rate of approximately 11.4%. An additional 25 companies started the survey but did not complete it. A comparison of incomplete and complete responses showed no systematic differences. Therefore, the incomplete datasets were excluded to ensure an unbiased and methodologically sound evaluation. With a response rate of 11.4 % for fully completed questionnaires, the survey performed at the higher end of typical online surveys, which generally yield lower rates than other formats such as telephone surveys and are often in the single-digit range [22,23]. Accordingly, the sample provides a representative overview of the German sawmill industry in terms of company size and production capacity and offers insights into current practices and challenges.

4. Results and Discussion

4.1 Characteristics of the Participating Sawmills

The 31 surveyed sawmills were classified according to their annual sawn timber production in cubic meters into small (<5,000 m³, 8 companies), medium (5,000–50,000 m³, 18 companies), and large enterprises (>50,000 m³, 5 companies). The number of employees ranged from fewer than 10 (35% of the companies) to 10–49 (48%), 50–250 (13%), and more than 250 employees (3%). The turnover distribution shows 19% of the companies with less than €1 million, 14 companies with €1–10 million, 4 companies with €10–50 million, and 1 company with more than €50 million; for 6 companies, no turnover information was provided. The composition of the participating companies reflects the overall structure of the industry, allowing inferences about sector-specific trends and characteristics (see Figure 3).

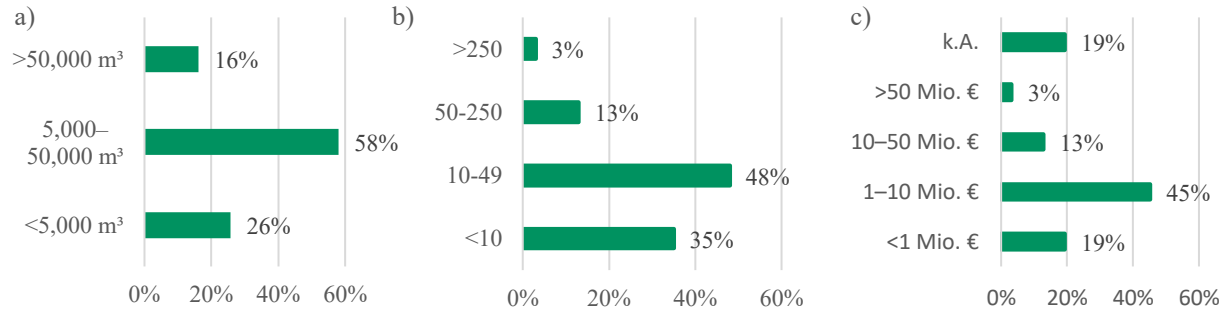


Figure 3: Classification of the surveyed sawmills by a) annual sawn timber, b) number of employees, c) turnover

Of the surveyed sawmills, 26 companies (84%) specialized in softwood and 5 companies (16%) in hardwood. The total sawn timber production of the participating sawmills amounted to 3,115,300 m³ and is distributed by wood type as follows: spruce, fir, and douglas fir 62%, pine and larch 35%, beech and other hardwoods 2%, and oak 2%. This largely corresponds to the classification of harvested wood reported by the Federal Statistical Office. [24] Deviations from Federal Statistical Office figures, especially for pine/larch and beech/other hardwoods, are likely due to the higher response rate from southern Germany.

4.2 Economic Aspects

To assess the economic situation and market dynamics, the survey collected information on both the revenue structure and the economic development from 2022 to 2024. On average, approximately 84% of revenue came from the sale of sawn timber, and 16% from by-products of the sawing process. Regarding the utilization of these by-products, 75% of the companies sold them as fuel, 72% to the wood-based panel industry, and 58% to agricultural customers. In addition, 50% of the companies used by-products for in-house energy generation, 14% for further processing into pellets, and 3% in other ways. No company left by-products unused.

Of the 31 participating sawmills, between 2023 and 2024 three companies recorded revenue growth, 15 companies experienced a decline, and seven maintained stable revenue levels. The average decline was approximately 11% and affected all company size categories. The three companies with revenue growth belonged to the group of small sawmills (<5,000 m³). Compared to 2022, the revenue decline was even more pronounced, with an average decrease of 22.7%. Only two companies were able to increase their revenue – by 2.3% and 33%, respectively – both also in the small sawmill category. Six companies did not provide revenue data.

Subsequently, the companies were asked to assess how six key economic challenges – energy costs, material availability, skilled labor shortages, price developments in the timber market, competitor pricing, and weakness in the construction sector – had affected their production costs over the past five years. The companies rated the impact as “none,” “moderate,” or “strong.” The results are shown in Figure 4.

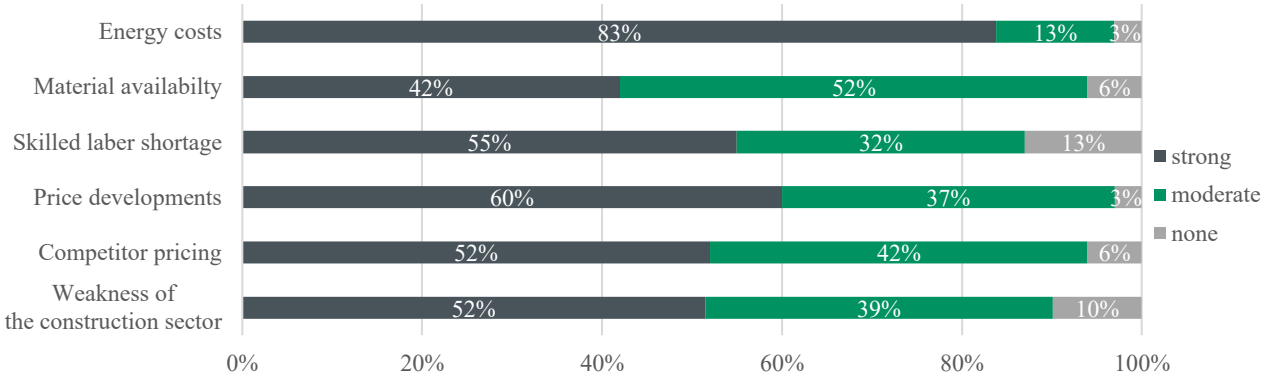


Figure 4: Assessment of the impact of economic challenges on production costs over the past five years.

The current economic situation was assessed as neutral by 61% of the surveyed companies, poor by 35%, and very poor by 3%. Differentiation by company size revealed no significant differences.

Regarding future market development, 16% of participants expect positive growth potential, 48% anticipate a stable development without significant changes, and 35% foresee a negative trend. The assessments vary by company size:

- Among small sawmills, sentiment is cautious, with 50% reporting neutral and 50% negative expectations,
- medium-sized companies show a mixed picture, with 17% positive, 50% neutral, and 33% negative assessments,
- large sawmills tend to be more optimistic, with 40% expecting positive growth, 40% stable development, and 20% negative expectations.

Overall, the results reflect a subdued but stable economic situation. Future expectations are predominantly cautious to neutral, with larger companies appearing somewhat more optimistic due to better capital and sales structures compared to smaller and medium-sized enterprises. Figure 5 illustrates the assessment of the current economic situation as well as the outlook for the coming years.

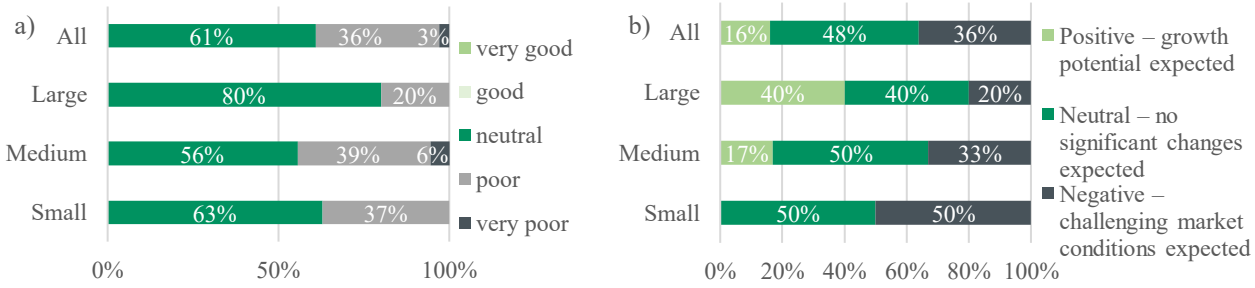


Figure 5: Assessment of the economic situation and future outlook by company size. a) Current economic situation; b) Expected market development.

4.3 Technological Aspects

Initially, the companies were asked about the sawing technologies used in their plants. Circular saws are employed by 65% of the surveyed sawmills, with 3% using them exclusively and the remainder in combination with frame or band saws. Pure frame saws are used by 19% of the companies, while 16% operate a combination of frame and band saws. The results reveal a heterogeneous machine landscape, with circular saws being the most commonly used technology.

The next step examined the current issues associated with the saw blades in use. The most frequently reported problems were rapid wear (41%) and poor cutting quality (38%). In addition, companies reported heat damage (13%), material breakage (9%), vibrations (6%), and high noise levels (6%). It should be noted that respondents were free to interpret “poor cutting quality” at their own discretion. In the context of sawmill

operations, this typically refers to surface roughness, uneven cuts, or deviations from intended dimensions. These issues reflect key challenges related to process stability, tool life, and energy input.

Building on this, participants were asked about overarching technological development needs. The greatest demand was identified in the areas of automation (81%) and energy efficiency (72%). Digitalization (56%) was also considered relevant, but ranked lower in priority. Only 6% of the companies reported no further technological development needs. These priorities are closely aligned with the previously identified issues related to tools and processes.

Regarding the specific further development of saw blades, several technological approaches were identified as particularly important. The highest priority was given to reducing energy consumption (68%) and increasing tool life through new coatings (58%). Improved running stability (45%) and self-sharpening saw blades (26%) were also considered relevant innovation directions. Reduced heat generation (19%), integrated sensors for wear detection (16%), and noise reduction (6%) were viewed as less critical. No additional aspects were reported in the open "Other" category. A comparison of the most common current operational problems and the identified development needs can be found in Figure 6.

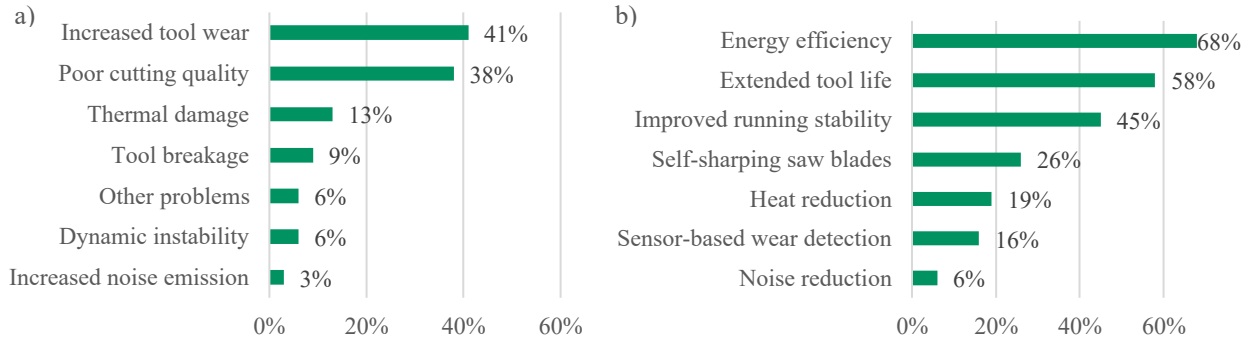


Figure 6: a) Most common operational problems, b) perceived development needs

Overall, it becomes clear that the technological challenges in the sawmill industry are closely linked to the identified development priorities. Issues such as high wear rates, insufficient cutting quality, or thermal stress are directly reflected in the demand for more energy-efficient, more stable-running, and longer-lasting tools. The analysis therefore confirms that there is a need for innovation along the entire process chain – from machine configuration and tool technologies to automated and energy-optimized process control.

4.4 Ecological Aspects

The participating companies were first asked about sustainability measures already implemented. The focus was on investments in renewable energy and resource efficiency. Around 71% of the companies operate photovoltaic systems, followed by the use of residual wood for energy generation (61%) and energy-efficient machinery (55%). Downstream processes for residual wood utilization (19%) and technical measures such as reduced kerf widths (13%) play a smaller role. Only 6% of the companies have not yet implemented any sustainability measures. Overall, the results indicate a strong willingness within the sector to integrate ecological improvements into operations.

The next step examined changes in wood quality and availability. The results indicate significant ecological shifts: 90% of respondents reported an increasing share of damaged wood, caused by drought, bark beetle infestations, and other climate-related stress factors. In addition, 68% observed a decline in certain tree species. Changes in wood properties, such as lower density or wider growth rings due to accelerated growth, were reported less frequently (10%). No company indicated that they had observed no changes – a clear indication that ecological developments are now having widespread impacts on the raw material base. These findings align with the German Forest Condition Report [25], highlighting that ecological changes are directly affecting sawmill operations.

With regard to the future availability of different tree species, most companies expect a moderate but noticeable change in species composition (67%). Around 30% even anticipate a very strong shift, in which completely new tree species could dominate the current composition. A largely stable development is hardly expected. These assessments point to the growing importance of climate-resilient tree species and, consequently, to long-term structural changes in the supply of timber.

Finally, the impact of forest conversion on operational practices was assessed. The most frequently cited consequence, reported by 48% of the companies, is a change in wood quality, which necessitates adjustments in the production process. Nineteen percent see new market opportunities through alternative tree species, while another 19% highlight the need for technical adaptations, for example to machines, frame saws, or saw blades. For 13%, forest conversion currently has no significant impact. Overall, it becomes evident that ecological change is not only altering the raw material base but is increasingly influencing strategic and technological decisions within the companies.

5. Conclusion and Outlook

This study provides a structured insight into current economic, technological, and ecological challenges in the German sawmill industry across different company sizes and identifies resulting innovation needs in sawing processes and related tool technologies. Data were collected via a standardized online survey, which was sent to 272 companies in the German sawmill industry. With 31 fully completed questionnaires, the response rate was 11.4%, which is above the average reported for industrial online surveys. Incomplete datasets were consistently excluded to ensure a methodologically sound and unbiased analysis.

The results of the study demonstrate that economic, technological, and ecological factors in sawmills are increasingly interrelated. The identified operational issues, in particular high tool wear, quality instabilities, and thermal stress, are directly reflected in the stated development needs. This indicates that existing process and tool technologies are reaching their performance limits, necessitating systematic further development. At the same time, assessments of wood availability and the effects of forest conversion suggest that the industry faces long-term structural adjustments. The increasing variability of wood quality imposes new demands on machines, processes, and tools, further highlighting the strategic importance of technological innovation.

From these findings, three primary areas for future research and technological development can be identified. First, there is a substantial need for durable and energy-efficient saw blades that can withstand the mechanical and thermal stresses imposed by increasingly heterogeneous raw materials while reducing overall energy consumption. This is particularly relevant for small sawmills, which represent the majority of operations and could contribute significantly to CO_2 reduction. At the same time, high investment costs highlight the need for cost-effective solutions accessible to smaller enterprises.

Second, the advancement of integrated digital monitoring systems is essential. Such systems enable continuous monitoring of production processes and tool wear, facilitating early detection of deviations, optimizing process stability, and supporting data-driven decisions.

Third, adaptive technologies for processing variable and heterogeneous wood assortments are increasingly critical. These technologies allow sawmills to respond flexibly to changes in wood quality and species composition, maintain product quality, minimize waste, and ensure economic viability under changing raw material conditions.

Overall, the study illustrates that the sawmill industry is undergoing a transformation in which efficiency improvement, process stability, and adaptability to ecological changes are central fields of action. The results provide a solid basis for systematically identifying future research priorities and for guiding technological innovations.

Acknowledgements

All topics presented in this paper are outcomes of the research project “Development of a high-performance circular saw blade with reduced heat generation – CryoCut” (Entwicklung eines Hochleistungskreissägeblattes mit reduzierter Wärmeaufnahme – CryoCut). This project is supported by the Federal Ministry of Food and Agriculture (BMEL) on the basis of a decision by the German Bundestag and is managed by the Fachagentur Nachwachsende Rohstoffe e.V. (FNR) under the funding program “Sustainable Renewable Resources” (Nachhaltige Erneuerbare Ressourcen – FPNR). The authors would like to express their gratitude for the funding that made this research work possible.

References

[1] Statistisches Bundesamt, 2025. Schadholzaufkommen nimmt weiter ab: Holzeinschlag 2024 um 13,3 % gesunken. Statistisches Bundesamt. https://www.destatis.de/DE/Presse/Pressemitteilungen/2025/04/PD25_144_41.html. Accessed 13 February 2026.

[2] Bundesministerium für Ernährung und Landwirtschaft, 2024. Der Wald in Deutschland: Ausgewählte Ergebnisse der vierten Bundeswaldinventur, Bonn. https://www.bmleh.de/SharedDocs/Downloads/DE/Broschueren/vierte-bundeswaldinventur.pdf?__blob=publicationFile&v=17. Accessed 13 February 2026.

[3] Bundesministerium für Landwirtschaft, Ernährung und Heimat, 2025. Holzmarkt - BMEL Statistikportal. https://www.bmel-statistik.de/forst-holz/holzmarkt. Accessed 13 February 2026.

[4] Statistisches Bundesamt (Destatis), 2025. Holzeinschlag: Deutschland, Jahre, Holzartengruppen: Code: 41261-0001. https://www-genesis.destatis.de/datenbank/online/statistic/41261/table/41261-0001. Accessed 13.02.206.

[5] Statistisches Bundesamt (Destatis), 2025. Erzeugerpreisindizes der Produkte des Holzeinschlags aus den Staatsforsten: Deutschland, Monate, Produkte des Holzeinschlags: Code: 61231-0002. https://www.destatis.de/DE/Themen/Wirtschaft/Preise/Landwirtschaftspreisindex-Forstwirtschaftspreisindex/Tabellen/Erzeugerpreise-Holzeinschlag.html#241824. Accessed 13 February 2026.

[6] Jauck, G., 2024. Säger massiv unter Druck. https://www.holzkurier.com/schnittholz/2024/04/saeger_massiv_unterdruck.html. Accessed 13 February 2026.

[7] Kiel, A., Kiel, U., 2023. Branchenanalyse Sägeindustrie: Im Spannungsfeld von Ökonomie und Umweltschutz. Forschungsförderung Working Paper 303, Düsseldorf.

[8] Bundesministerium für Landwirtschaft, Ernährung und Heimat, 2025. Holzmarktbericht 2024: Abschlussergebnisse für die Forst- und Holzwirtschaft des Wirtschaftsjahres 2024, Bonn. https://www.bmleh.de/SharedDocs/Downloads/DE/Broschueren/holzmarktbericht-2024.pdf?__blob=publicationFile&v=3. Accessed 13 February 2026.

[9] Statistisches Bundesamt (Destatis), 2025. Rohholz, Holzhalbwaren: Deutschland, Jahre, Rohholzart, Holzerzeugnisse, Bestandsveränderungen: Code: 42341-0001. https://www-genesis.destatis.de/datenbank/online/statistic/42341/table/42341-0001/search/s/cm9oaG9seg%3D%3D. Accessed 13 February 2026.

[10] Döring, P., Gieseking, L., Mantau, U., 2020. Sägeindustrie 2018: Einschnitt- und Produktions- volumen, Hamburg. https://infro.eu/img/pdf/rohstoffmonitoring/R%202020%2001%20S%C3%A4geindustrie.pdf. Accessed 13 February 2026.

[11] Mantau, U. (Ed.), 2023. Holzrohstoffbilanzierung - Kreislaufwirtschaft und Kaskadennutzung: 20 Jahre Rohstoffmonitoring Holz. Fachagentur Nachwachsende Rohstoffe e.V. (FNR), Gülzow-Prüzen, 04 pp.

[12] Statistisches Bundesamt (Destatis), 2025. Beschäftigte und Umsatz der Betriebe im Verarbeitenden Gewerbe: Deutschland, Jahre, Wirtschaftszweige (WZ2008 2-4-Steller Hierarchie): Code: 42271-0003. https://www-genesis.destatis.de/datenbank/online/table/42271-0003/table-toolbar/search/s/NDIyNzEtMDAwMw==#filter=JTdCJTIyaGlkZUVtcHR5Q29scyUyMiUzQWZhbHNlJTJDJTIyaGlkZUVtcHR5Um93cyUyMiUzQWZhbHNlJTJDJTIyY2FwdGlvbiUyMiUzQSU1QiU3QiUyMnZhcmlhYmxlSWQlMjIlM0ElMj. Accessed 13 February 2026.

[13] Fachagentur Nachwachsende Rohstoffe e. V. (FNR), 2018. Rohstoffmonitoring Holz: Erwartungen und Möglichkeiten, Gülzow-Prüzen. https://www.charta-fuer-holz.de/fileadmin/charta-fuer-holz/dateien/service/mediathek/Broschuere_Rohstoffmonitoring_Holz_Web_neu.pdf. Accessed 13 February 2026.

[14] Koskivaara, A., Lähtinen, K., Toppinen, A., 2025. Longitudinal analysis of financial performance of sawmills and wood construction element producers in Finland. Silva Fenn 59 (2).

[15] Makkonen, M., 2019. Renewing the sawmill industry: studies on innovation, customer value and digitalization. Diss. For. 2019 (269).

[16] Ranacher, L., Stocker, R., Riegler, M., Schwarzbauer, P., Hesser, F., 2023. Digitalization of small and medium sized sawmills in Austria: a survey about business processes. Eur. J. Wood Prod. 81 (1), 267–280.

[17] Mensah, P., Pimenta, A.S., Melo, R.R. de, Amponsah, J., Tuo, G., Chakurah, I., Ampadu, S.D., Buckman, I., Nikoi, M., Minkah, E., Miranda, N.d.O., Medeiros, P.L. de, 2025. The Global Supply Chain of Wood Products: A Literature Review. Forests 16 (7), 1036.

[18] Gottlöber, C., 2023. Zerspanung von Holz und Holzwerkstoffen: Grundlagen - Systematik - Modellierung - Prozessgestaltung, 2., aktualisierte Auflage ed. Hanser, München, 1226 pp.

[19] Wagenführ, A., Scholz, F. (Eds.), 2018. Taschenbuch der Holztechnik, 3., aktualisierte Auflage ed. Fachbuchverlag Leipzig im Carl Hanser Verlag, München, 1567 pp.

[20] Magerhans, A., 2016. Marktforschung: Eine praxisorientierte Einführung. Springer Gabler, Wiesbaden, 360 pp.

[21] Homburg, C., 2020. Marketingmanagement: Strategie - Instrumente - Umsetzung - Unternehmensführung, 7., überarbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden, Heidelberg, 1473 pp.

[22] Daikeler, J., Bošnjak, M., Lozar Manfreda, K., 2020. Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. Journal of Survey Statistics and Methodology 8 (3), 513–539.

[23] Kreis, H., Wildner, R., Kuß, A., 2024. Marktforschung. Springer Fachmedien Wiesbaden, Wiesbaden.

[24] Statistisches Bundesamt (Destatis), 2025. Gesamteinschlag nach Holzartengruppen. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Wald-Holz/Tabellen/gesamteinschlag-holzartengruppen.html.

[25] Bundesministerium für Landwirtschaft, Ernährung und Heimat, 2025. Ergebnisse der Waldzustandserhebung 2024, Bonn. https://www.bmleh.de/SharedDocs/Downloads/DE/Broschueren/waldzustandserhebung-2024.pdf?__blob=publicationFile&v=9. Accessed 13 February 2026.

Biography

M.Sc. Dominik Roppelt (*1989) joined the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth as a researcher in 2022. He holds a Master's degree in Industrial Engineering and Mechanical Engineering from OVGU Magdeburg, with a specialization in production engineering. His research focuses on the development and optimization of circular saw blades and associated sawing technologies, with an emphasis on efficient and sustainable production methods.

Dr.-Ing. Julian Schmidt (*1993) joined the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth as a scientific assistant in 2019. He holds a PhD in Engineering Science from the University of Bayreuth. Since 2022, he has been Senior Engineer and head of the sustainable production department.

Dr.-Ing. Benjamin Thorenz (*1987) holds a PhD in Engineering Science and has been a researcher at the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth. Since 2020 he has been Senior Engineer of the department and Head of the Keylab Cutting.

Prof. Dr.-Ing. Frank Döpper (*1970) holds the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth and is division director Manufacturing and Process Engineering at Fraunhofer IPA. For more than twenty years, he worked in technology and investment management and as a site and production manager before he was appointed professor at the University of Bayreuth.

8th Conference on Production Systems and Logistics

Towards Predictive Quality Control In Large-Scale Battery Cell Finalization: A Pillar-Based Framework

Julian Wonneberger[1], Patric Keister[1], Jonathan Krauß[1], Simon Lux[1,2] and Louis Blumenschein[1]

[1]*Fraunhofer Research Institution for Battery Cell Production FFB, Bergiusstr. 8, 48165 Münster, Germany*
[2]*University of Münster, Institute of Business Administration at the Department of Chemistry and Pharmacy, Leonardo Campus 1, 48149 Münster, Germany*

Abstract

Cell finalization, including End-of-Line (EoL) testing and quality control (QC), constitutes a critical bottleneck in lithium-ion battery manufacturing. These processes dominate production time and impose substantial demands on energy, capital investment, and factory floor space. Despite their high resource intensity, prevailing QC methods fail to reliably assure long-term performance or effectively mitigate cell-to-cell variability, making QC a decisive cost and risk factor. This paper gives insight into common QC practices in battery cell finalization, derives key limitations, and discusses recent approaches for improving cell characterization, with a focus on predictive quality control (PQC). An overview of recent advances in PQC research is given, distinguishing between approaches for immediate quality assessment and cycle life prediction. To bridge the gap between academic research and industrial application, four essential pillars for industrially viable PQC systems are derived from industrial QC practices, big data quality management, and standardization efforts: *Functionality, Scalability, Integrability* and *Adaptability*. These pillars form a research agenda aimed at implementing PQC in large-scale battery manufacturing and advancing QC capabilities toward industrial deployment.

Keywords

Battery Cell Production; Battery Cell Finalization; Predictive Quality Control; Machine Learning; Long-Term Cell Performance

1. Introduction

The global demand for lithium-ion batteries (LIBs) is rapidly increasing, driven by the electrification of transportation and the growing need for energy storage technologies in the context of the energy transition [1–3]. While LIBs are already widely used in consumer electronics and are becoming essential for renewable energy integration and electric vehicles, their production processes remain economically challenging in Europe. Depending on the specific literature, the material costs are stated to make up approx. 60-80 % of the total costs. The remainder of the costs almost completely consists of the production costs [4–7]. Manufacturing LIB cells is a complex, multi-stage process comprising electrode production, cell assembly, and cell finalization [4,8]. The resource demand of each phase varies depending on the specific production setup and factory location, resulting in substantial differences in their contribution to total manufacturing costs [9,4,10,11]. Independent of factory design, scrap remains costly throughout all stages due to its cumulative impact and should therefore be minimized [12].

Cell finalization encompasses the final production steps, typically including soaking, formation, degassing, self-discharge testing and EoL testing. QC is integrated at or after these steps. While initial electrolyte filling is not always classified as part of cell finalization, additional filling operations may occur during this phase.

DOI: https://doi.org/10.15488/20974
ISSN: 2701-6277

After electrolyte injection, cells undergo soaking, usually at elevated temperatures, to ensure complete wetting of electrode and separator pores [13,14]. Electrolyte filling and soaking apply only to cells with liquid electrolytes. Formation constitutes the first electrochemical cycling of the cell. Its primary purpose is the controlled formation of passivation layers, the solid-electrolyte interphase (SEI) on the anode and the cathode-electrolyte interphase (CEI) on the cathode, which stabilize electrode-electrolyte interfaces and significantly reduce long-term degradation [13,10,14]. Gas evolution during formation affects performance and evolving gases should therefore be removed during degassing. Depending on cell design, degassing may occur after formation, between formation sub-steps, or continuously throughout formation. After the final degassing, cells are hermetically sealed to protect internal components from contamination [13,14]. To ensure safety and verify performance, additional QC procedures are carried out during EoL testing, which typically assess key battery parameters multiple times for varying test configurations [13,14]. As literature inconsistently uses *EoL* to refer either to dimensional cell control or to QC activities throughout the finalization stages, we consistently refer to these procedures as *quality control* to avoid ambiguity. The sequence of process steps in finalization differs across factories, formats and chemistries, with some manufacturers revisiting partially completed steps. Reported cost shares for cell finalization and its sub-steps vary significantly; formation alone may account for 25-45 % of total manufacturing cost [9,4,14]. It is acknowledged that finalization requires substantial time; typically, 2-3 weeks. This duration is dominated by the self-discharge test (aging), which often spans more than a week, while preceding and subsequent operations can still require several days in total [13–15].

The following chapter discusses the state of the art in battery production QC, including challenges and approaches for improvement. While PQC demonstrates high potential, a gap persists between academic research and industrial application. Chapter 3 reviews the current state of PQC research, Chapter 4 proposes four foundational pillars to bridge this gap, and Chapter 5 provides conclusions and an outlook.

2. Quality Control in Battery Cell Finalization

QC in battery cell finalization is central to ensuring safety, validating performance, and protecting downstream system reliability. QC procedures typically quantify electrochemical and physical parameters under multiple test configurations [13,14]. Key electrochemical metrics include discharge capacity and energy, coulombic efficiency (CE) with emphasis on initial CE (ICE), and internal resistance characterized under direct current (DCIR) and alternating current (ACIR) stimuli. Complementary checks encompass dimensions, weight, leakage, and defect identification via self-discharge screening [16].

2.1 State of the Art Quality Control

In industrial practice, cycling tests derive discharge capacity, discharge energy, and (I)CE, with values highly sensitive to measurement parameters (e.g., C-rate, temperature, voltage windows) and differing between initial SEI/CEI formation cycles and subsequent stabilized cycles. Slower C-rates improve data fidelity but extend test durations, increase energy consumption through thermal conditioning, and raise capital requirements for formation/testing infrastructure [13,14]. Impedance measurements determine ACIR and, via electrochemical impedance spectroscopy (EIS), resolve process-specific contributions across frequency. However, production QC often limits itself to ACIR checks at one or a few frequencies due to equipment and time constraints. Pulse power tests for DCIR apply brief constant-current pulses at a defined State of Charge (SoC), with the voltage response serving as a proxy for operational polarization. Self-discharge tests screen for latent mechanical defects by monitoring voltage decay over days to weeks; elevated-temperature variants shorten duration but increase energy use and thermal infrastructure needs [13,16,14].

Operationally, testing outcomes categorize cells as scrap or saleable, with saleable cells graded based on their individual performance according to predefined thresholds [13,14]. These thresholds are based on development series testing following company-specific and/or international guidelines, e.g. the IEC 62660

standard. Development series testing includes comprehensive safety verification, including cycle life and abuse tests. Since these tests aren't applicable to sold cells, statistical sampling based on standards such as IEC 62133 assesses population level safety and validates manufacturing processes, leaving gaps in individual cell characterization.

2.2 Challenges in Conventional Quality Control

Despite the significant efforts, current QC predominantly certifies immediate performance and safety at factory exit and often falls short of predicting long-term performance. As a result, commercial cells within a single batch show divergent degradation rates [17]. Moreover, variant degradation behavior has been observed both within and between cell grades [18]. These findings underscore that existing QC methods fail to capture cell-to-cell variations that either amplify or emerge during operational use. The root of this predictive failure runs deeper than initially apparent. Initial cell-to-cell variations in capacity and impedance do not reliably correlate with subsequent degradation or performance divergence, indicating that standard QC and factory cell matching are insufficient to ensure long-term uniformity within battery packs [17,19]. Similarly, manufacturer quality grades do not consistently reflect true cell performance, with variation and outliers present within nominally identical batches [18]. Understanding the sources of variation is key to addressing this challenge. Cell-to-cell variations can originate across the entire battery lifecycle, from raw material inconsistencies and manufacturing processes to module assembly, shipping, and operation [20]. For battery cell QC, variations that are detectable during manufacturing are of particular importance, as their efficient detection and prediction not only optimize QC systems but also enable downstream adaptive usage.

To frame the quality problem systematically, ATTIA ET AL. distinguish two dimensions: *defects* and *non-conformity* [21]. Defects are defined as internal cell flaws that can cause malfunctions and safety incidents; many remain latent and only manifest after prolonged cycling. Non-conformity refers to deviations from intended cell design that do not necessarily cause immediate failures (also termed cell-to-cell variation). While non-conformities may lie within accepted tolerances, they signal inadequate process control and elevate the risk of latent defects. Both dimensions can have severe consequences at module or pack level due to interaction effects in series/parallel configurations and thermal-electrical coupling.

These fundamental limitations motivate a shift toward QC strategies that are both more resource-efficient and have a stronger focus on long-term performance. Battery cell finalization QC thus faces a dual challenge: reducing resource consumption while enhancing predictive power regarding long-term behavior.

2.3 Advances in Quality Control

Multiple approaches exist to address quality issues in cell finalization via measurement techniques, spanning destructive and non-destructive methods across ex-situ, in-situ, and operando modalities [22]. Inline-applicable, non-destructive options include electrochemical techniques (e.g., rapid EIS, high power pulse characterization (HPPC), galvanostatic intermittent titration technique (GITT)), physically coupled sensing (temperature, pressure, thickness/expansion), image-based inspection (2D X-ray/CT, optical/IR thermography), and mechanically induced methods (ultrasonic testing) [20,13,23]. Method selection must balance cost, throughput, technology readiness level (TRL), sensitivity, environmental compatibility, and data actionability. Among these approaches, electrochemical QC represents a challenge due to its condition-dependent variability and direct impact on performance and longevity. However, recent studies found electrochemical data well suited to detect cell-to-cell inhomogeneities [20]. Given that electrochemical measurements are already routinely performed during finalization, utilization of this data offers a particularly powerful and cost-effective lever for advancing QC.

Advanced data analysis represents a key enabler for this leverage [21]. For this purpose, three primary approaches exist, each with distinct characteristics [24,25]. *Data-driven approaches*, including statistical analysis, machine learning (ML), and deep learning (DL), are highly effective for analyzing diverse, high-dimensional data. They offer great flexibility, enable rapid inference, and require less domain expertise in

battery characterization than physics-based approaches, but typically rely on large datasets. *Physics-based* approaches, such as equivalent circuit models and electrochemical simulations, interpret data through mechanistic principles. They provide deep insights and can operate with limited data but demand extensive expert knowledge and careful parameterization, especially during cell finalization, when cell characteristics are still evolving. Depending on the model complexity, they are often computationally intensive, which can hinder real-time deployment. *Hybrid approaches* integrate the two paradigms to balance their strengths. This integration can be *external* (e.g., ML-based parameter identification for physics models, physics-based feature extraction for ML) or *internal* (e.g., physics-informed machine learning). These methods aim to reduce data requirements, enhance model generalizability, and maintain interpretability, while requiring moderate domain knowledge and computational resources [24,26,27]. Overall, there is a trade-off among data requirements, domain expertise, inference speed, and generalization ability.

Applying advanced analytical approaches to finalization data enables PQC, which offers substantial benefits for cell manufacturing. PQC can significantly reduce the need for lengthy cycle life testing while maintaining accuracy, and early identification of non-conforming cells prevents wasted resources, lowering costs and environmental impact [28,29]. Reliable grading enables intelligent grouping at module and pack level, mitigating risks from cell-to-cell variability, preventing weak-cell chain reactions, and enabling adaptive operational strategies that extend system lifetime [17]. PQC enables scalable quality assessment by using electrochemical signatures as surrogate health indicators, thereby defining target variables that link defects to process parameters and support targeted tolerance tuning and proactive optimization [21,30,31]. Collectively, these capabilities underscore the strategic value of implementing advanced QC.

3. Recent Research on Predictive Quality Control in Battery Cell Finalization

While battery cell quality represents a well-established concern, research has predominantly focused on condition and remaining lifetime prediction (i.e. state of health (SoH), remaining useful life (RUL)) during operational use rather than manufacturing-stage quality assessment, largely due to limited public availability of production data [25]. This chapter provides an overview of recent research on PQC in battery cell manufacturing. It summarizes key approaches for predicting both immediate cell performance (Chapter 3.1) and long-term cell performance (Chapter 3.2) based on production data (see Figure 1).

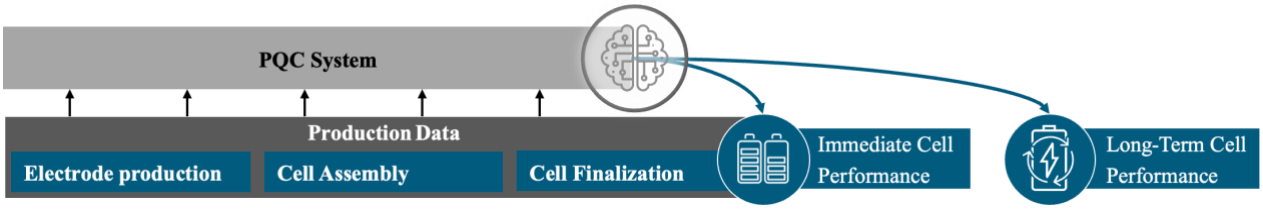


Figure 1: PQC systems use production data to predict the immediate or long-term cell performance.

3.1 Related Work on Predictive Quality Control for Immediate Cell Performance

Existing research can be grouped according to the data used for quality prediction: data from post-formation quality tests, formation data, and upstream production data. The sooner a battery or its components are identified as scrap, the fewer resources are wasted. However, the further an intermediate product is from the finalized battery, the more difficult it is to link it to final quality. Moreover, high-quality components do not guarantee a high-quality cell [32–34,29].

YUAN ET AL. developed a hybrid method to increase the speed of determining a battery's usable capacity after completing the formation process. The group used two batches of 5,000 prismatic lithium iron phosphate (LFP) battery cells from a pilot production line for their study. They used a neural network for analyzing features from the capacity test's beginning parts to predict the result of the capacity test. They were able to achieve a root-mean-square error (RMSE) of 0.197 Ah (0.18 %) for the capacity prediction,

reduce the capacity grading process's energy consumption by approx. 37 %, and its duration by approx. 50 %. One major challenge that was identified was generalizing the model between the two batches [29]. LIU ET AL. continued the concept of YUAN ET AL. by adding formation features to an XGBoost-based model, further shortening the capacity grading step. Their PQC system reduced the grading step's duration by approx. 80 % and the utilized energy by approx. 56 % compared to the unmodified process. The model was originally developed using 1,000 graphite/NCM-based Li-ion cells. To demonstrate the transferability of their concept, they also graded 2,886 graphite/LFP-based cells with a reported mean absolute percentage error (MAPE) of 0.2456 % [33]. LI ET AL. developed an XGBoost-based voting ensemble enabling capacity-grading-free PQC using only features from electrolyte filling and formation reaching an MAPE of 0.4979 %. A second model that additionally incorporates early capacity-grading data reduced MAPE down to 0.3501 %. Both models were trained on production data from 171,666 cells manufactured on an industrial line outputting 2,000–10,000 cells per day over 30 days [32]. NIRI ET AL. examined correlations between electrode physical properties and half-cell capacities using 133 pilot-line cells with NMC622 cathodes with unstated anode. Their ML model achieved RMSEs of approx. 3 % for cathode half-cell capacities at various C-rates and showed robustness to added noise but could not predict full-cell capacities [34].

In summary, recent work shows that hybrid PQC schemes combining feature extraction and ML can markedly accelerate or even bypass capacity grading, yielding time and energy savings of up to 80 % and 56 % by progressively leveraging data from partial capacity tests, formation, and upstream production steps.

3.2 Related Work on Predictive Quality Control for Long-Term Cell Performance

Recent studies that predict long-term performance of battery cells focus on using data from finalization and electrolyte filling.

STOCK ET AL. developed a hybrid framework for early quality classification and cycle life prediction of lithium-ion cells using production data from soaking (EIS data), formation, and early cycling on 29 NMC111/graphite pouch cells with intentionally varied electrolyte fill levels. They engineered 24 features and compared elastic-net regression with neural networks. The best model achieved a 10.1 % test error for cycle life regression using 29 inputs, including the first 20 cycles. For binary lifetime classification (threshold 250 cycles), the best model reached 97 % accuracy using data including the first 20 cycles. They found a decreased accuracy including the first five cycles compared to using only formation data [28]. ZOU ET AL. developed a hybrid framework for battery quality classification and lifetime prediction using formation data. The approach extracts over 100 features from voltage, current, capacity, and energy signals, combining statistical descriptors, differential analyses and electrochemical indicators (consumed lithium, relaxation voltages). After feature selection and dimensionality reduction they apply k-nearest neighbors for quality classification, and ridge regression for lifetime targets (EOL80/EOL70, capacity retention at 300/350/400 cycles, knee point). Applied to the dataset of 40 NMC/graphite pouch cells from [35] (two formation protocols, baseline and fast; aging at 25 °C and 45 °C). The framework achieved 89.74 %/89.7 % classification accuracy, mean percentage errors of 6.50 %/5.45 % for EOL80 prediction, and retention-at-300-cycles errors as low as 2.74 % (25 °C) and 1.05 % (45 °C) [31]. RHYU ET AL. propose a formation-only hybrid framework for lithium-ion cycle-life prediction. Starting from Q(V) of the last discharge, fused lasso identifies informative voltage windows and yields two simple difference features in the 3.57–3.66 V range, which are used with a lightweight regressor (e.g., Random Forest). Evaluated with nested cross-validation on 178 SC-NMC532/graphite pouch cells spanning 62 formation protocols, the resulting two-feature model achieves an average RMSE of approx. 98 cycles [25]. YANG ET AL. proposed a purely data-driven DL framework for prediction of lithium-ion battery cycle life using formation data. Their approach transforms sequential measurements (voltage, current and capacity) from the second cycle into images using a sliding window method and applies convolutional neural networks (CNN) for automatic feature extraction and prediction. Evaluated on the dataset created in [36] (124 LFP/graphite cells), the model achieved an RMSE of 76.81 cycles [37].

In summary, recent studies demonstrate that production data from finalization enables cycle life prediction and quality classification. Both purely data-driven and hybrid approaches that combine physics-based features with machine learning have proven effective.

3.3 Challenges for Predictive Quality Control of Battery Cell Performance

Building on the limitations of conventional QC discussed in Section 2.2, PQC introduces additional challenges. Materials, cell chemistry and finalization protocols strongly influence the availability and characteristics of extractable features, thereby limiting the transferability of PQ-models [13,14,35]. Together with the pronounced cell-to-cell variability across batches and manufacturers [18,20,17,19], this leads to limited generalization, even when sizable datasets are available [29].

For immediate cell performance prediction, studies report large datasets, but the underlying data is usually proprietary and not publicly accessible [32,33,29]. For long-term cell performance, existing approaches have typically been demonstrated only on small-scale or pilot lines, with restricted sample sizes and heterogeneous data management practices [25,28,37,31]. As a result, high-quality, large-scale datasets that consistently link production conditions, early-stage signatures and verified long-term outcomes are scarce. This hampers systematic validation of PQC models and makes it difficult to prove that detected patterns correspond to practically relevant defects or non-conformities.

To date, no publications describe the integration of PQC approaches into IT systems at industrial scale. This represents a critical barrier to practical implementation, as seamless integration into manufacturing execution systems (MES), quality management systems (QMS), supervisory control and data acquisition (SCADA) and data infrastructure is essential for real-time decision-making and continuous production optimization [30,38]. Missing concepts for end-to-end traceability, standardized interfaces and structured feedback of PQC outputs into continuous-improvement cycles further limit the impact of current research [21,29]. Moreover, most methods have not yet been demonstrated under gigafactory-scale conditions, where prediction for hundreds of thousands of cells per day, heterogeneous sensor setups and tight latency constraints impose additional scalability requirements [21,32].

To bridge the gap between current research and industrial implementation, the following chapter introduces four key pillars that structure these challenges into requirements for industrially applicable PQC systems.

4. Pillars for Predictive Quality Control Systems in Battery Cell Production

As highlighted in previous chapters, PQC has the potential to transform battery cell finalization by enabling faster and more reliable quality decisions, thereby addressing several of the challenges discussed in Section 2.2. As outlined in Section 3.3, however, moving from academic prototypes to industry-grade systems requires more than high predictive accuracy: in regulated manufacturing environments, PQC must be robust, explainable, verifiable, secure, and aligned with existing quality management systems and regulatory constraints. Persistent challenges such as cell-to-cell and batch-to-batch variation further increase the demands on any quality control strategy, including PQC.

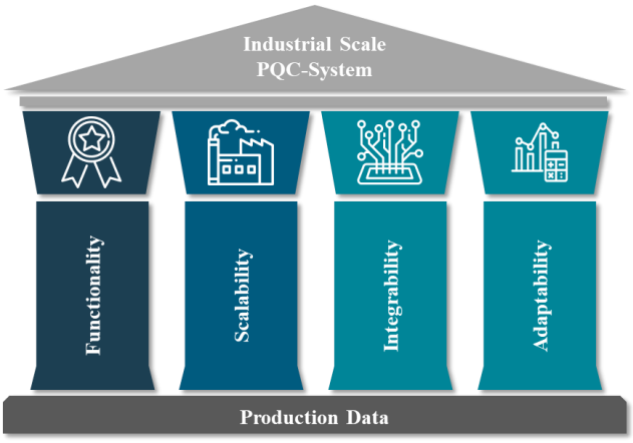


Figure 2: The proposed pillars include Functionality, Scalability, Integrability and Adaptability.

Traditional standards for quality assurance, metrology, and functional safety (e.g., ISO 9001, IATF 16949, VDA volumes, IEC standards) define overarching requirements and structural guidelines for industrial quality control, while battery-specific standards such as IEC 62660 and IEC 61960 specify detailed QC procedures with defined measurement uncertainties.

Together, they enable robust QC in battery production. More recently, AI-focused standards such as ISO/IEC TR 5469 and ISO/IEC TR 24028 have begun to address the development and governance of AI systems, particularly regarding trustworthiness and security. Yet, no standard currently specifies how established QC procedures in industry can be replaced or complemented by predictive quality control, as existing standards cover QC for measured characteristics but do not define how to monitor and assure the quality of predicted characteristics when they are used to replace conventional measurements in industrial battery production. Additionally, PQC relies on high-frequency, heterogeneous process data, data-driven models, and tightly integrated IT/OT architectures, and thus introduces requirements that are only weakly reflected in current standards: handling model drift and in-field retraining, propagating measurement uncertainty into predictions, versioning models and data pipelines, ensuring deterministic inference, and providing explainable, actionable outputs for process optimization. To systematically address these gaps, this chapter introduces four essential pillars that any industrial PQC system for battery cell production should fulfill (see Fig. 2): Functionality, Scalability, Integrability, and Adaptability. Derived by mapping the limitations of current QC and PQC (Sections 2.2 and 3.3) against established quality and AI standards and by incorporating practical experience and expert input, these pillars provide a conceptual framework for translating predictive methods into scalable, trustworthy, and compliant solutions for battery manufacturing.

Functionality describes the technical validity of PQC predictions along the dimensions: predictive performance and uncertainty, robustness and determinism, metrological validation, and explainability with actionable confidence information. The primary function of a Predictive Quality Control (PQC) system is to accurately predict the quality parameters of finalized battery cells, whereby the uncertainty of the system's input data must remain below the desired tolerance of the predicted parameter, as measurement errors propagate through the prediction, which is particularly critical for data-driven models that can turn small input uncertainties into disproportionately large prediction errors [39,40]. The system must remain robust under a defined level of input noise and behave deterministically, providing identical predictions for identical input data [41,42]. Further, predictions generated by the PQC system must undergo regular verification, analogous to the calibration of physical measurement instruments, by comparing them against results from electrochemical or physical characterization tests [39,40]. To ensure that predictions are made for the right reasons, the system should provide explainable outputs and associated confidence levels, enabling informed decisions on whether cells require additional electrochemical testing [41–43]. ***Scalability*** captures the ability of PQC systems to handle gigafactory-level throughput data streams while remaining sensor-independent and data-efficient. In gigafactories, the production of hundreds of thousands to millions of battery cells per day is a typical output [21]. Consequently, the PQC system should be able to handle high-frequency and high-volume process data. Due to the parallelized production infrastructure in gigafactories, measurements of the same parameter must be harmonized across sensors and lines so that PQC predictions remain independent of the specific measurement device used. To maximize the impact of PQC on resource efficiency, data acquisition and processing should focus on parameters that are critical for process control and quality prediction while keeping measurement and storage overhead low [41,42,32]. ***Integrability*** describes how PQC can be embedded into existing production IT while preserving process awareness and cell-level traceability. To integrate production data and enable feedback of quality predictions, a PQC system must ensure seamless integration with production control systems such as MES, enterprise resource planning (ERP) and supervisory control and data acquisition (SCADA) to facilitate coordinated data flow across processes [30,38]. Standardized interfaces help use data correctly and avoid hidden maintenance [30,44]. Given that data is highly dependent on the finalization recipe, PQC systems, including data interfaces and data models, should be tightly linked to recipe management [14]. Reliable PQC operation further depends on unambiguous mapping of input data and predictions to the correct cell, process step and cell state (e.g. temperature, SoC), preserving end-to-end traceability and synchronization with the physical process. ***Adaptability*** captures the capability of PQC to remain reliable under gradual process drift and deliberate process changes through continuous monitoring, governed retraining and ongoing validation. Variations in

the process naturally lead to cell-to-cell variations, and additional changes to the battery production process are likely, for example when switching material suppliers or adjusting cell finalization protocols. To adapt reliably to such changes and variations, a structured monitoring plan is essential [45]. Even minor shifts in the data can directly affect model predictions; therefore, both data streams and model outputs must be continuously monitored [44]. A retraining strategy with clear triggers and comprehensive data/model versioning should be established for the PQC system, including rollback capability. This requires systematic data storage, as well as continuous validation to ensure sustained functionality.

While *Functionality* forms the technical prerequisite, its practical value is only realized through *Scalability* and *Integrability*, and its long-term viability is secured by *Adaptability*. Together, these pillars establish a comprehensive framework for implementing an industrially viable PQC system.

5. Conclusion and Outlook

This paper examined challenges of current QC in battery cell finalization. Conventional QC is resource-intensive, time-consuming, and yet fails to reliably account for long-term performance or effectively mitigate cell-to-cell variability. PQC leveraging production data represents a particularly promising pathway to reduce finalization bottlenecks while improving reliability. Recent studies demonstrate that PQC approaches for immediate cell performance have achieved significant time reductions and energy savings in capacity grading. Approaches for predicting long-term performance have demonstrated high accuracy in early cycle life prediction and quality classification using only formation and initial cycling data. Both purely data-driven methods and hybrid approaches combining physics-based features with ML have proven effective.

However, transitioning from academic prototypes to industrial deployment requires addressing fundamental challenges. To bridge this gap, this work proposed four essential pillars: *Functionality, Scalability, Integrability, and Adaptability,* that define requirements for industrially viable PQC systems. These pillars extend beyond predictive accuracy to encompass robustness, explainability, verifiability, IT integration, and compliance with quality management systems and regulatory frameworks.

Future research should build on the proposed pillars and pursue a coordinated research roadmap. First, the pillars should be translated into measurable criteria and metrics, thereby enabling systematic evaluation, comparison and, ultimately, certification of PQC systems. Quantifiable metrics would allow existing work on PQC in battery cell finalization to be reassessed beyond predictive accuracy, considering criteria such as robustness, explainability and data efficiency. This step would further help to identify gaps and potential extensions of the proposed pillar framework. To further address *Scalability and Integrability,* standardized PQC test protocols and data schemas should be developed, including harmonized definitions of features, labels and traceability information. This would support reproducible experiments, fair benchmarking of PQC across cell formats and finalization protocols, and foster the creation of open, high-quality datasets. With respect to *Adaptability*, research is needed on model- and data-centric approaches for monitoring and managing drift in industrial PQC deployments, including retraining triggers aligned with existing quality and safety standards. Systematic integration of upstream process data should be advanced to leverage early performance indicators, maximizing PQC value across design, ramp-up, and continuous improvement.

While this paper establishes a conceptual foundation towards implementing PQC in large-scale battery production, continued interdisciplinary collaboration between academia, industry, and standardization bodies is essential to realize its full potential. Successfully implemented, PQC systems can fundamentally transform battery cell finalization by reducing resource consumption, improving QC, and enabling data-driven continuous improvement, thereby advancing the economic viability and sustainability of battery production.

Acknowledgments

This work was supported by the project FOOD (EFRE-20800230), funded by the European Union and the state of North Rhine-Westphalia, Germany, under the EFRE/JTF Programme NRW 2021–2027. Furthermore, the authors acknowledge funding from the project "FoFeBat – Forschungsfertigung Batteriezelle Deutschland", supported by the Federal Ministry of Research, Technology and Space under the reference numbers 03XP0256, 03XP0416, 03XP0501A, and 03XP0590A.

References

[1] Da Deng, 2015. Li-ion batteries: basics, progress, and challenges. Energy Science & Engineering 3 (5), 385–418.

[2] Fleischmann, J., Hanicke, M., Horetsky, E., Ibrahim, D., Jautelat, S., Linder, M., Schaufuss, P., Torscht, L., Alexandre van de Rijt, 2023. Battery 2030: Resilient, sustainable, and circular. McKinsey & Company, January 16.

[3] Sterner, M., Stadler, I., 2017. Energiespeicher - Bedarf, Technologien, Integration. Springer Berlin Heidelberg, Berlin, Heidelberg, 876 pp.

[4] Duffner, F., Mauler, L., Wentker, M., Leker, J., Winter, M., 2021. Large-scale automotive battery cell manufacturing: Analyzing strategic and operational effects on manufacturing costs. International Journal of Production Economics 232, 107982.

[5] Kwade, A., Haselrieder, W., Leithoff, R., Modlinger, A., Dietrich, F., Droeder, K., 2018. Current status and challenges for automotive battery production technologies. Nat Energy 3 (4), 290–300.

[6] Soldan Cattani, N., Noronha, E., Schmied, J., Frieges, M., Heimes, H., Kampker, A., 2024. Comparative Cost Modeling of Battery Cell Formats and Chemistries on a Large Production Scale. Batteries 10 (7), 252.

[7] Wentker, M., Greenwood, M., Leker, J., 2019. A Bottom-Up Approach to Lithium-Ion Battery Cost Modeling with a Focus on Cathode Active Materials. Energies 12 (3), 504.

[8] Kampker, A., 2014. Elektromobilproduktion. Springer Berlin Heidelberg, Berlin, Heidelberg, 290 pp.

[9] Degen, F., Krätzig, O., 2024. Modeling Large-Scale Manufacturing of Lithium-Ion Battery Cells: Impact of New Technologies on Production Economics. IEEE Trans. Eng. Manage. 71, 6753–6769.

[10] Liu, Y., Zhang, R., Wang, J., Wang, Y., 2021. Current and future lithium-ion battery manufacturing. iScience 24 (4), 102332.

[11] Orangi, S., Strømman, A.H., 2022. A Techno-Economic Model for Benchmarking the Production Cost of Lithium-Ion Battery Cells. Batteries 8 (8), 83.

[12] Örüm Aydin, A., Zajonz, F., Günther, T., Dermenci, K., Berecibar, M., Urrutia, L., 2023. Lithium-Ion Battery Manufacturing: Industrial View on Processing Challenges, Possible Solutions and Recent Advances. Batteries 9 (11), 555.

[13] Kampker, A., Heimes, H., Offermanns, C., Wennemar, S., Robben, T., Lackner, N., 2023. Optimizing the Cell Finishing Process: An Overview of Steps, Technologies, and Trends. WEVJ 14 (4), 96.

[14] Schomburg, F., Heidrich, B., Wennemar, S., Drees, R., Roth, T., Kurrat, M., Heimes, H., Jossen, A., Winter, M., Cheong, J.Y., Röder, F., 2024. Lithium-ion battery cell formation: status and future directions towards a knowledge-based process design. Energy Environ. Sci. 17 (8), 2686–2733.

[15] Wood, D.L., Li, J., An, S.J., 2019. Formation Challenges of Lithium-Ion Battery Manufacturing. Joule 3 (12), 2884–2888.

[16] Lee, J., Yang, M.-J., Lee, D., Ha, S., Doo, S., Lee, S., Shim, J.-H., 2025. Voltage drop screening for defective cells during the battery formation process. Journal of Energy Storage 131, 117579.

[17] Devie, A., Baure, G., Dubarry, M., 2018. Intrinsic Variability in the Degradation of a Batch of Commercial 18650 Lithium-Ion Cells. Energies 11 (5), 1031.

[18] Ank, M., Kröger, T., Schreiber, M., Lienkamp, M., 2023. Experimental analysis of lithium-ion cell procurement: Quality differences, correlations, and importance of cell characterization. Journal of Energy Storage 66, 107430.

[19] Rumpf, K., Naumann, M., Jossen, A., 2017. Experimental investigation of parametric cell-to-cell variation and correlation based on 1100 commercial lithium-ion cells. Journal of Energy Storage 14, 224–243.

[20] Beck, D., Dechent, P., Junker, M., Sauer, D.U., Dubarry, M., 2021. Inhomogeneities and Cell-to-Cell Variations in Lithium-Ion Batteries, a Review. Energies 14 (11).

[21] Attia, P.M., Moch, E., Herring, P.K., 2025. Challenges and opportunities for high-quality battery production at scale. Nature Communications 16 (1), 611.

[22] Nature Communications, 2022. Improving the fundamental understanding of batteries via operando measurements. Nature Communications 13 (1), 4723.

[23] McGovern, M.E., Bruder, D.D., Huemiller, E.D., Rinker, T.J., Bracey, J.T., Sekol, R.C., Abell, J.A., 2023. A review of research needs in nondestructive evaluation for quality verification in electric vehicle lithium-ion battery cell manufacturing. Journal of Power Sources 561, 232742.

[24] Borah, M., Wang, Q., Moura, S., Sauer, D.U., Li, W., 2024. Synergizing physics and machine learning for advanced battery management. Communications Engineering 3 (1), 134.

[25] Rhyu, J., Schaeffer, J., Li, M.L., Cui, X., Chueh, W.C., Bazant, M.Z., Braatz, R.D., 2025. Systematic feature design for cycle life prediction of lithium-ion batteries during formation. Joule 9 (5), 101884.

[26] Li, W., Zhang, J., Ringbeck, F., Jöst, D., Zhang, L., Wei, Z., Sauer, D.U., 2021. Physics-informed neural networks for electrode-level state estimation in lithium-ion batteries. Journal of Power Sources 506, 230034.

[27] Singh, S., Ebongue, Y.E., Rezaei, S., Birke, K.P., 2023. Hybrid Modeling of Lithium-Ion Battery: Physics-Informed Neural Network for Battery State Estimation. Batteries 9 (6), 301.

[28] Stock, S., Pohlmann, S., Günter, F.J., Hille, L., Hagemeister, J., Reinhart, G., 2022. Early Quality Classification and Prediction of Battery Cycle Life in Production Using Machine Learning. Journal of Energy Storage 50, 104144.

[29] Yuan, Y., Kong, X., Hua, J., Pan, Y., Sun, Y., Han, X., Yang, H., Li, Y., Liu, X., Zhou, X., Lu, L., Wang, H., Ouyang, M., 2023. Fast grading method based on data driven capacity prediction for high-efficient lithium-ion battery manufacturing. Journal of Energy Storage 73, 109143.

[30] Kampker, A., Heimes, H., Lingohr, P., Schmied, J., Clever, H., Dorn, B., 2024. Developing a concept for the implementation of predictive quality in battery production. Procedia CIRP 126, 823–828.

[31] Zou, J., Gao, Y., Frieges, M.H., Börner, M.F., Kampker, A., Li, W., 2024. Machine learning for battery quality classification and lifetime prediction using formation data. Energy and AI 18, 100451.

[32] Li, H., Kong, X., Ruan, G., Yang, K., Zhou, J., Chen, C., 2025. Fast capacity grading and capacity grading-free process for lithium-ion battery production lines. Journal of Energy Storage 136, 118397.

[33] Liu, Z., Xu, R., Wang, H., 2025. Capacity prediction method of lithium-ion battery in production process based on eXtreme Gradient Boosting. Ionics 31 (2), 1759–1777.

[34] Niri, M.F., Liu, K., Apachitei, G., Ramirez, L.R., Lain, M., Widanage, D., Marco, J., 2021. Machine learning for optimised and clean Li-ion battery manufacturing: Revealing the dependency between electrode and cell characteristics. Journal of Cleaner Production 324, 129272.

[35] Weng, A., Mohtat, P., Attia, P.M., Sulzer, V., Lee, S., Less, G., Stefanopoulou, A., 2021. Predicting the impact of formation protocols on battery lifetime immediately after manufacturing. Joule 5 (11), 2971–2992.

[36] Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., Fraggedakis, D., Bazant, M.Z., Harris, S.J., Chueh, W.C., Braatz, R.D., 2019. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 4 (5), 383–391.

[37] Yang, W., Min, F., Xie, J., Yang, H., 2025. Ultra-Early Prediction of Lithium-Ion Battery Cycle Life Based on Visualized Single-Cycle Data. IEEE Trans. Power Electron. (40), 7342–7353.

[38] Mohring, L., Jaspers, W., Schmetz, A., Roth, D., Hamacher, N.C., Kampker, A., 2024. Digital Battery Cell Manufacturing Systems: Approaching a scalable and versatile IT Architecture Design. Procedia CIRP 130, 492–497.

[39] DIN Deutsches Institut für Normung e. V. DIN EN ISO 9001:2015-11.

[40] Verband der Automobilindustrie (VDA), 2021. VDA 5 Mess- und Prüfprozesse.

[41] International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC), 2020. ISO/IEC TR 24028:2020.

[42] International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC), 2024. ISO/IEC TR 5469:2024.

[43] Zhang, R., Ji, C., Zhou, X., Liu, T., Jin, G., Pan, Z., Liu, Y., 2024. Capacity estimation of lithium-ion batteries with uncertainty quantification based on temporal convolutional network and Gaussian process regression. Energy 297, 131154.

[44] Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., Dennison, D., 2015. Hidden Technical Debt in Machine Learning Systems.

[45] Chatterjee, A., Ahmed, B.S., Hallin, E., Engman, A., 2022. Quality Assurance in MLOps Setting: An Industrial Perspective, 8 pp. https://arxiv.org/pdf/2211.12706. Accessed 4 December 2025.

Biography

Julian Wonneberger is a research associate in the Department of Digitalization at the Fraunhofer Research Institution for Battery Cell Production (FFB) and a PhD candidate at RWTH Aachen University. He holds an MSc in mechanical engineering from TH Köln – University of Applied Sciences. His research focuses on enabling end-to-end traceability and developing data-driven methods for process optimization in industrial battery cell production, with particular emphasis on cell finalization.

Patric Steven Keister is a research associate at the Fraunhofer Research Institution for Battery Cell Production FFB and a PhD candidate at the University of Münster. He obtained an MSc at RWTH Aachen University in 2023. Afterwards, he worked for research institutes as a research associate. Since joining Prof. Lux's group in 2025, his research focuses on quality control in the finalization stage of industrial battery cell production.

Dr.-Ing. Jonathan Krauß leads the digitalization of battery cell production at Fraunhofer FFB. He holds a doctorate in engineering from RWTH Aachen University and is a recognized expert in the digital transformation of industrial manufacturing. Dr. Krauß has authored numerous scientific publications and has served as chairman at international conferences. He has successfully bridged research and industry in over 50 publicly and privately funded projects.

Prof. Dr. Simon Lux is Professor of Applied Electrochemical Energy Storage Technology and Economic Chemistry at the University of Münster and a member of the Institute Management of the Fraunhofer Research Institution for Battery Cell Production FFB as well as the Executive Board of the "FoFeBat" project.

Louis Blumenschein is a research associate at Fraunhofer Research Institution for Battery Cell Production FFB specializing in cell finalization. He obtained an MSc at Karlsruhe Institute of Technology in 2023. At Fraunhofer FFB, he focuses primarily on battery modeling, specifically model-based process optimization.

8th Conference on Production Systems and Logistics

Intelligent Process Control In Paperboard Compression-Drawing Using Digital Twins And Machine Learning

Muhammad Faisal Yaqoob[1], Klara Liesegang[1], Lena Berthold[2], Stephan Kronenberger[1], Christer-Clifford Schenke[1], Steffen Ihlenfeldt[1,3]

[1]*Fraunhofer Institute for Machine Tools and Forming Technology, Germany*
[2]*TUD Dresden University of Technology, Chair of Processing Machines / Processing Technology, Germany*
[3]*TUD Dresden University of Technology, Chair of Machine Tools Development and Adaptive Controls, Germany*

Abstract

Disposable plastic packaging is widespread due to its low cost, efficiency, and ease of production. On the contrary, its slow biodegradation contributes to long-term environmental pollution. This favors paperboard as a more sustainable and recyclable alternative. However, to effectively control the compression-drawing process for producing paperboard cups, it is essential to manage the interactions between the material and various process parameters (e.g., blankholder force, tool temperature), as well as ambient conditions. Targeted moistening of the material has been shown to stabilize the process, further increasing the overall complexity. The main challenge is coordinating all these interdependent parameters during each production cycle to ensure a stable paperboard production process. As part of a joint project between Fraunhofer IWU and TU Dresden, an intelligent process control system was developed and implemented on a paperboard cup compression-drawing demonstrator. This paper highlights the development and deployment of this control system, emphasizing the integration of digital twin technology and machine learning to adaptively regulate process parameters of paperboard compression-drawing. The result of such control system is a robust manufacturing process and consistent cup quality, paving the way for more intelligent and sustainable packaging processes.

Keywords

Digital Twins; Machine Learning; Intelligent Process Control; Compression-drawing, Paperboard

1. Introduction and motivation

The packaging industry constitutes a globally significant economic sector. It's projected to increase from $1.14 trillion in 2024 to $1.38 trillion by 2029, representing a 3.89% compound annual growth rate over the forecast period. (2024–2029) [1]. And while single-use plastics in packaging are increasingly restricted due to pollution and non-recyclability, paperboard offers a sustainable and recyclable alternative, leading many countries to prioritize eco-friendly packaging solutions. The Packaging Act [2] enacted in Germany is an example of such to promote circular economy and increasing recycling rates. Replacing plastic requires production rates comparable to conventional packaging. Although the compression-drawing process can meet this demand, drawing paperboard remains a challenging despite extensive research work. In [3] the interactions between paperboard material and compression-drawing tool were extensively evaluated. This research established that the quality of paperboard parts produced by compression-drawing can be primarily measured by number of wrinkles. High wrinkle numbers correlate with high dimensional quality and stability. In [4] and [5], the influences and interactions of dominant process parameters, such as blankholder force, friction, etc. were thoroughly investigated. Among these, die temperature and blankholder force were

DOI: https://doi.org/10.15488/20975
ISSN: 2701-6277

identified as the most significant parameters, with both improving wrinkle distribution and producing finer, more uniform patterns. In [6,7], It was found that ambient conditions such as humidity and temperature alter the paper's moisture content, influencing how the paper deforms and how defects develop during drawing. The research from [8] further showed that targeted moistening can stabilize paperboard cups, with optimal formability achieved at 7–11% moisture. This moisture range softens the material, reducing defects and ensuring shape stability under controlled temperatures.

The challenge therefore lies in ensuring robust process performance in paperboard compression-drawing process by coordinating targeted material moistening with the identified key process parameters (blankholder force and die temperature) for each production cycle. This highlights the relevance of developing an intelligent process control system that could enable continuous adaptation of process parameters across production cycles. Here, Machine Learning (ML) algorithms may be applied for adaptive control under varying process and environmental conditions. A sheet metal drawing process described in [9] utilized a data-driven controller approach, which is analogous to the methodology of the current research. The authors combined digital twins and machine learning involving key parameters (local lubrication, friction, and blankholder force) to predict the produced part geometry in real time, achieving improved part quality post-optimization. While this approach has been successfully demonstrated for sheet metal drawing application, its effectiveness in paperboard drawing remains to be established. Consequently, the presented paper investigates the development and deployment of an intelligent process control system that predicts the quality of a drawn paperboard cup based on moisture, die temperature, and blankholder force, and optimizes these parameters throughout each production cycle on an industrial control system. This research work was done through a joint project between Fraunhofer IWU and TUD on a paperboard cup production demonstrator. An intelligent control system, named the Technology Control Model (TCM), was realized by combining digital twin approach with ML.

The paper is organized as follows: Chapter 2 reviews the literature concerning the primary technologies integrated into TCM, specifically digital twins and ML, while Chapter 3 details the used research demonstrator and it's process flow. Chapter 4 presents the methodology for developing and deploying the TCM. Finally, Chapter 5 provides a summary of the findings and highlights the future outlook.

2. Digital twins and machine learning in production

This chapter introduces the topic of digital twins and machine learning, examining their distinct use cases within the production environment. Furthermore, establishing the relevance of these technologies to the TCM developed for the paperboard compression-drawing application.

2.1 Simulation models and Digital twin

Digital twin technology, a core Industry 4.0 innovation, models, simulates, and optimizes manufacturing systems, improving efficiency, reducing downtime, and supporting lifecycle management [10]. Due to the broad application of this technology across different sectors, there have been diverse range of definitions identified in a systematic review in [11] based on aspects like models, data, links, and functions. In [12] the digital twin was defined as a digital representation of a physical object or a system that reflects its properties, condition and behavior in real time. In [13] and [14], often the term digital twin was used for a simple CAD, a simulation model or a data based counterpart mirroring physical entity in real time. In [15], the digital twin concept was explored through interoperability and standardization, with the Asset Administration Shell (AAS) as the core architecture for managing and storing asset data throughout its lifecycle. According to the definition in [12], the digital twin is the final result of the development process, which begins with a simple 3D representation and proceeds via a simulation model. An example scenario of different digital twin applications across the three phases (planning, commissioning, and operation) is shown in Figure 1.

In Phase 1, the machine is not available and remains in the planning, concept, or ideation stage. The aim is to transition from real to digital, converting sketches into digital formats such as CAD data to support efficient development decisions.

In Phase 2 VC is defined as method for the early development and validation of programmable logic controller (PLC) code using a simulation model. In such simulation, kinematic structures are modelled onto available CAD data using joints and motors [16], [17] and [18]. When coupling such simulation with controls, there are mainly three configurations:

- Model in the Loop [MiL]: Simulation models are tested in a virtual simulation without physical hardware
- Software in the Loop [SiL]: Real software is integrated into a virtual simulation through emulated versions, but no physical hardware is involved.
- Hardware in the Loop [HiL]: Physical hardware is integrated into a virtual simulation to test real-time system interactions

Finally in Phase 3, the machine is operational and real-time data is available. The digital twin can now mirror the real machine, enabling data-based solutions such as anomaly detection, predictive maintenance, or edge solutions.

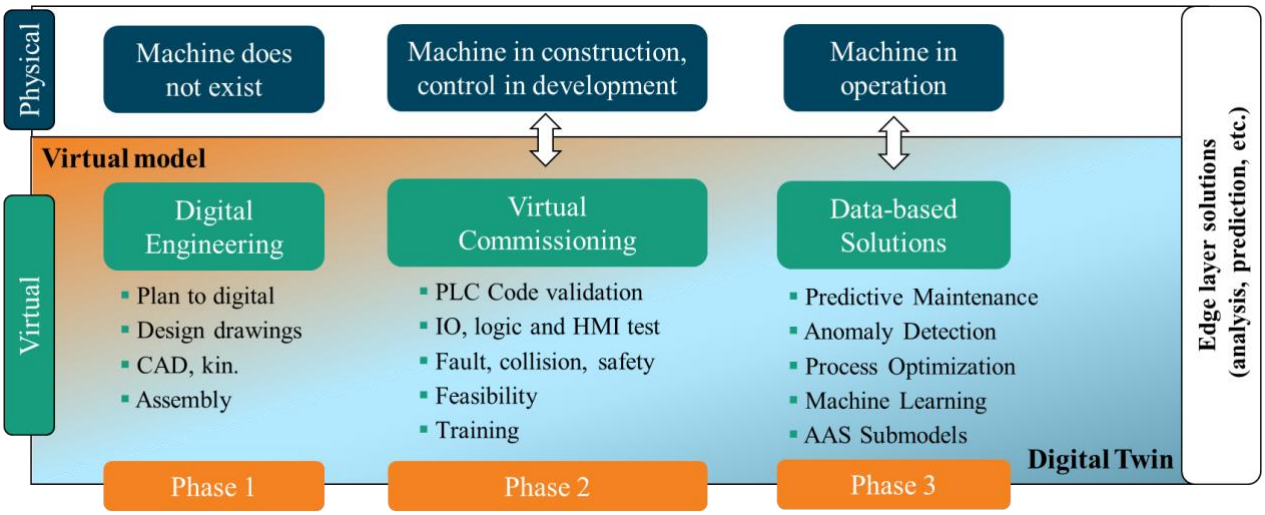


Figure 1: Phases of digital twins (from planning to operation)

During the digital engineering phase, the FE-Simulations facilitates the synthetic data generation required to train the ML model. Subsequently, in the VC stage, a simulation model is utilized for TCM validation prior to physical deployment. Consequently, in the third stage of the process, TCM is deployed into controls and functions as the required data-based solution to effectively control the process parameters.

2.2 Machine learning (ML)

Machine Learning (ML) refers to the field of Artificial Intelligence that allows applications to learn from data, instead of being programmed for every task, enabling the development of smart, real-time engineering applications [19]. Frequently used ML models in manufacturing, including Convolutional Neural Networks (CNN), Restricted Boltzmann Machines, Autoencoders, and Recurrent Neural Networks, must be carefully selected based on their use case [20]. In [21], a complex welding process was optimized via ML involving several process parameters like tool life, defects, process optimization, corrosion rate prediction, and grain size estimation. Different ML-models were studied where CNN achieved the highest accuracy of > 95% accuracy in image-based defect detection but required large datasets. This resulted in optimized parameters and good quality welds. In [22], 3D printing process was optimized via different ML approaches involving process parameters like bed temperature, cooling, layer speed and width, etc. Random Forest showed highest accuracy of 97.5%, suggesting it correctly predicts or classifies 97.5% of the instances in the test dataset.

Similarly, in [23], two optimization methods using ML were studied to intelligently determine the injection molding process parameters. Results showed that it's possible to intelligently control the process and achieve near to zero defects in final product, however more data collection was suggested. Depending on the type of problem, computing requirements and timeframe, experts chose the suitable ML-model. For example, tabular/structured data may go better with Random Forest while CNN is better suited to image processing.

Based on the literature studied, it is evident that digital twins and ML are known to be effective parts of industrial automation. The presented research combines both technologies into a Technology Controls Model (TCM) to enable robust paperboard production, aligning its manufacturability with that of plastic cup production. TCM is a data-driven controller trained with simulation data. It is based on three main components (ML model, quality assessment and optimization) and predicts the number and distribution of wrinkles based on the current condition of the paper and planned process parameters.

3. Paperboard cup production machine as project demonstrator

The paperboard cup production demonstrator utilized in this research project initially comprised three core modules: cutting, compression-drawing, and sealing. To address the needed parameter spectrum and to enable a part evaluation, the demonstrator was extended with two additional modules: humidification and quality assessment. The humidification module applies controlled moisture to the paperboard before the compression-drawing process, operating in sync with the system cycle, with adjustable duration and intensity. The quality module evaluates cup quality by detecting wrinkles using transmitted-light imaging with an industrial camera and surface light positioned after the compression-drawing process. Figure 2 illustrates the process flow (P1-P12) of the demonstrator machine used in this research work, showing the sequence of operations.

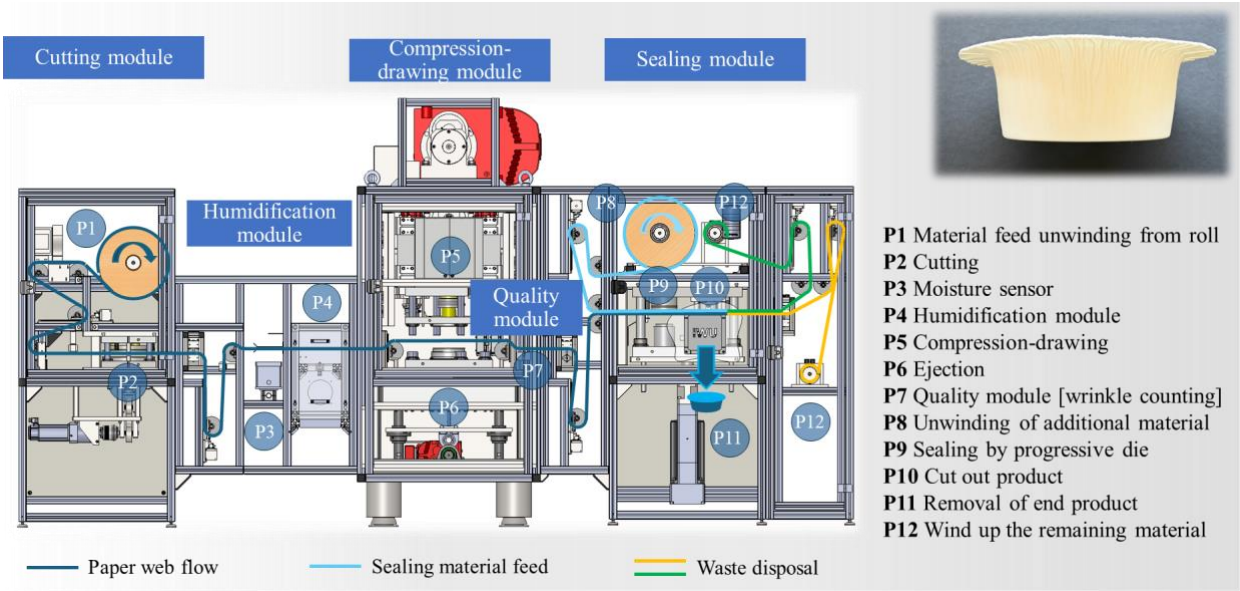


Figure 2: Process flow of demonstrator machine and produced paperboard cup

Operational control supports manual and automatic modes, with production setpoints defined via traditional input or a Python interface for a Python-based learning controller. The demonstrator is capable of producing drawn paperboard cups from a single web of material. During production, the target values, process variables, and quality parameters are recorded.

4. Methodology

This chapter introduces the methodology to develop and deploy the TCM. As shown in Figure 3, the method is divided into the development of the four modules: data acquisition, the development of the TCM, virtual commissioning and the deployment in the machine control as briefly detailed further:

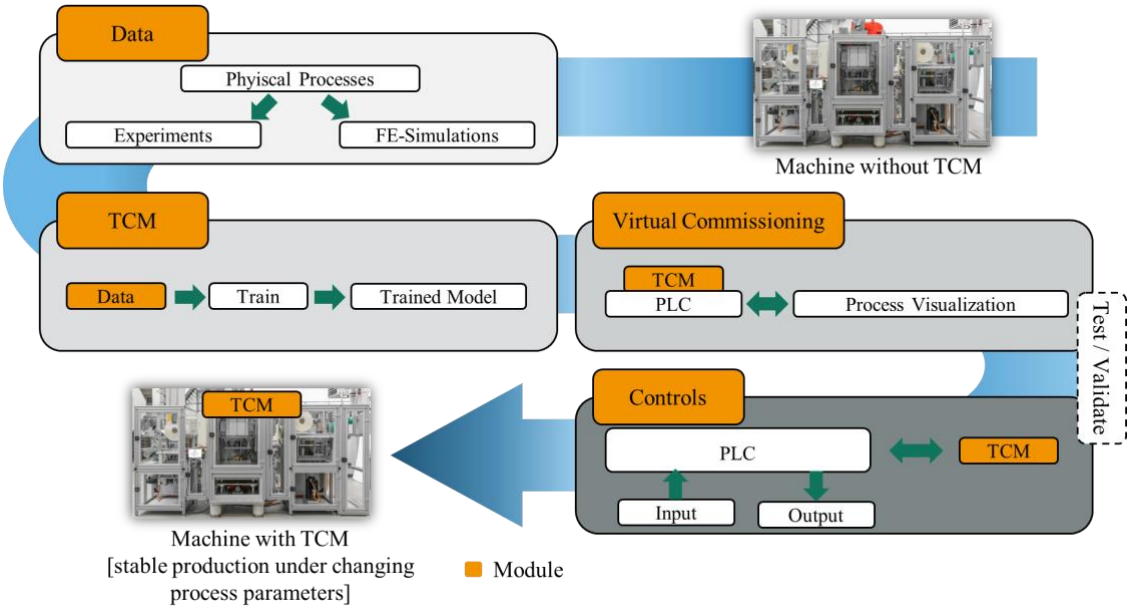


Figure 3: Methodology for TCM development and deployment

4.1 Data acquisition

The first step is to generate the synthetic dataset for the training of the TCM. This was achieved using a Finite-Element (FE) simulation which predicts cup geometry, especially the wrinkle pattern, depending on the process parameters and material properties. The FE-simulation was also employed to ascertain whether material failure occurs. It is noteworthy that in this study, the qualitative relationship between cup geometry and process parameters/ material conditions was sufficient. Consequently, relatively simple material and friction models were employed in order to reduce calibration effort and computing times. Tensile and friction tests were performed for varying material temperatures and moisture contents in order to calibrate the FE-simulation. Further, compression-drawing experiments were conducted on a laboratory setup with the same tool geometry as the production demonstrator. This experimental dataset was used to validate the FE-simulation in terms of material failure and wrinkle pattern for varying moisture content, tool temperature and blank holder force. Exemplary for the gathered data, Figure 4 presents the process window, which was analyzed as part of this research work. Further details on this analysis can be found in [24].

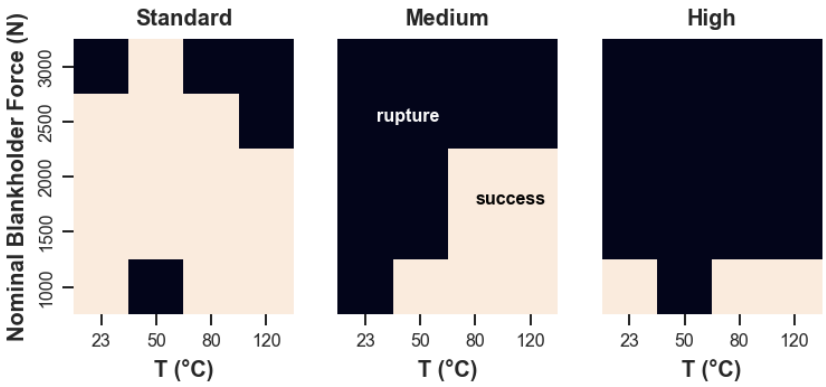


Figure 4: Process window for compression-drawing cups with conditioned material at three moisturization levels (Standard – Medium - High) [24]

Additionally, a comprehensive examination of the FE-simulation, also conducted as part of this research, is available in [25]. The validated FE-simulation reliably maps cup quality and process limits. However, due to the high computational intensity, its application for real-time predictions is not feasible. Therefore, as a second step, the TCM was developed.

4.2 Development of TCM

The TCM, which consists of three main components, operates according to the following scheme: the moisture content of each blank is measured at the production demonstrator prior to compression-drawing. Subsequently, an ML-model predicts the quality of the resulting cup. In the event of material failure, adjustment of the process parameters is imperative. Conversely, if an intact part is predicted, a quality algorithm assesses the wrinkle pattern (number and distribution) of the forecast. The objective is to fabricate parts that exhibit numerous, uniformly dispersed wrinkles. In the event that deficiencies are identified, it is necessary that the process parameters undergo adjustment. This alteration was facilitated by an optimization algorithm. The adjusted process parameters were communicated to the control of the production demonstrator and modified prior to the occurrence of a defective cup. Ultimately, the TCM facilitates an adaptive stroke-to-stroke control, thereby ensuring the consistent production of cups of high quality, even in the presence of changing material properties, such as fluctuating material moisture content.

4.2.1 Training of ML-Model

The ML-model was trained using the synthetic dataset from section 4.1. In the context of a stroke-to-stroke control, the parameters and material properties that are subject to variation during the process are of particular relevance. Preliminary findings as already discussed in Chapter 1 indicate that the moisture content of the blank, the blank holder force, and the temperature of the die are factors that must be taken into consideration. These three process parameters are, therefore, the inputs of the ML-model. The moisture content of the blank was calculated from the measured moisture content and the planned moistening duration. According to the inputs provided, the part geometry and wrinkle pattern were predicted. A convolutional neural network (CNN) was used for this purpose. Comprehensive details concerning the selection of the model, the specific architecture, and the outputs can be found in [25].

4.2.2 Quality Assessment and Optimization Algorithm

The predictions of the ML-model undergo an evaluation of the wrinkle pattern. For this purpose, the positions of nodes defined on the flange of the part serve as a signal. The evaluation of this signal achieved through the implementation of the discrete Fourier transform and yields a scalar value. This scalar value serves to not only quantify the number of wrinkles but also to assess their distribution. A thorough exposition of the quality algorithm can be found in [21]. In the event that the quality algorithm indicates substandard quality or material failure, the determination of optimized process parameters is executed via a grid search. More complex optimization strategies, such as Particle Swarm Optimization or Metaheuristics, require greater computational resources and are most advantageous when dealing with large-scale problems. In the current context, the optimization encompasses a mere three parameters, which are constrained within a limited range due to the constricted nature of the process window. Consequently, a grid search exhibits an advantageous effort-use ratio. The grid search varies the three inputs of the ML-model systematically within the process limits. For each parameter combination, a prediction is generated using the ML model and evaluated by the quality algorithm. Within the parameter combinations that do not exhibit material failure, the one with the highest quality forecast is automatically selected. The calculation of the requisite spraying duration is contingent upon the optimal material moisture content, and the adjusted process parameters are communicated to the control of the production demonstrator.

4.3 Virtual Commissioning (VC)

The primary objective of VC was to minimize safety risks and ensure a validated TCM. The initial plan involved connecting a virtualized version of demonstrator controller (Bosch-CML75) to the VC model to conduct SiL simulations involving "what-if" disturbance scenarios with generated virtual I/O's. However, since the necessary connector for the controller was unavailable, real-time communication was adopted as the alternative. Consequently, a functional integration check of the TCM was only possible within a virtual environment using Model-in-the-Loop (MiL) simulation connected to real-time controller values via OPC-UA. This ensured that the TCM responded to machine control values and operated continuously throughout the production cycles. The TCM was integrated inside the VC model in the Open Neural Network Exchange format (ONNX). ONNX provides an open-source framework designed to facilitate the interoperability of machine learning models across different platforms. As for the VC model development, CAD model of the paperboard cup demonstrator was imported, segmented into assemblies, and kinematized. C# logic was developed to manage the process flow and trigger the execution of the ONNX model. Finally, reference workpieces representing produced cups at various stages were integrated to illustrate material flow and achieve visualization of near-realistic process.

The resulting VC model, illustrated in Figure 5, incorporates the following technical features:

- Realistic simulation of the entire process flow
- Workpiece creation for visualization reference
- Model-in-the-Loop (MiL) simulation
- ONNX test in VC model
- OPC UA communication for data exchange (digital shadow)
- Can be used for employee training (integrated with VR)

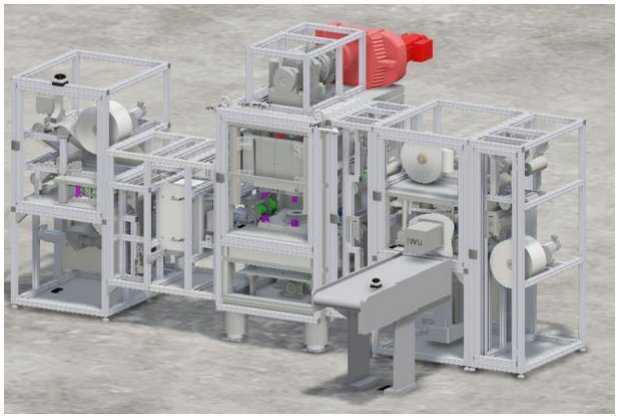


Figure 5: Virtual commissioning model of paperboard cup production demonstrator

4.4 Controls deployment

In this module, the TCM was deployed to the controls of the demonstrator machine with setup involving different applications integrated into the Industrial-PC (IPC), as illustrated in Figure 6 .

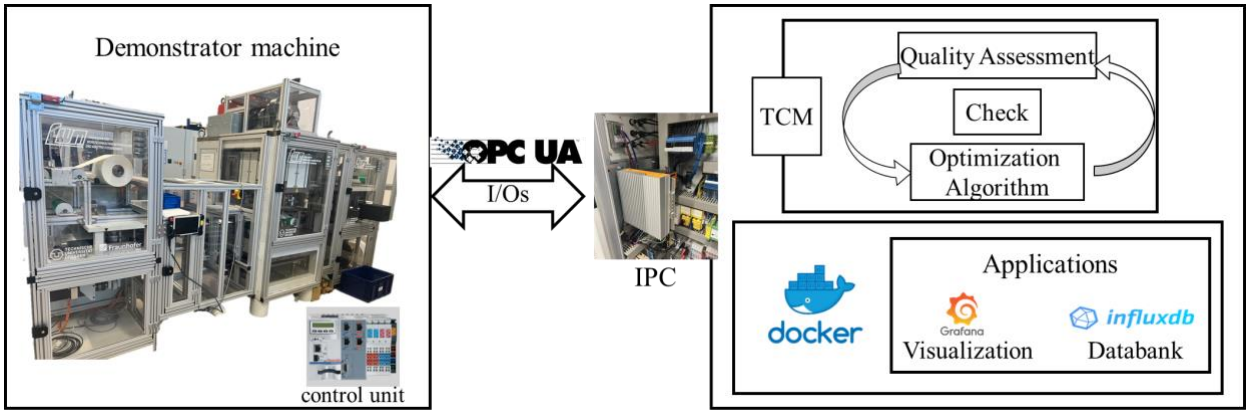


Figure 6: TCM deployment to demonstrator machine

The Bosch Rexroth CML 75 control unit, utilizing Sercos III for high-speed, real-time communication, forms the backbone of the current control system. This unit was programmed and configured using the IndraWorks Engineering Software, which facilitates the setup, commissioning, and diagnostics of the Bosch control

hardware. For data management, InfluxDB, a time-series database was configured to handle the high-frequency data streams generated by the machine. The InfluxDB instances were visualized through Grafana, a web-based tool that allows operators to monitor real-time data and visualize historical trends. Data exchange between the system components was managed using OPC UA (Unified Architecture), a standardized communication protocol that ensures secure, platform-independent data transfer across the network. The Bosch control unit utilized an integrated OPC UA server to communicate with a custom client on the IPC. The TCM module, Grafana, and InfluxDB were deployed via Docker on the IPC, which was directly connected to the demonstrator. The TCM operated on an IPC near the machine, communicated via OPC-UA and calculated a quality forecast for every production cycle. This ensured that wrinkle distribution remained within predefined tolerances. Finally, TCM performance was assessed based on part quality and adaptive control of temperature and humidification, by comparing its capabilities to conventional fixed setpoint specifications.

5. Results and Discussion

To evaluate how the TCM performs compared with conventionally fixed parameter setups, cycled production tests were performed. Under stable production conditions, fixed parameter setups achieved a 100% yield across 10 samples with an average wrinkle count of 95±5. In comparison, the TCM reached a 96% yield over a 50-cup cycle at a production rate of 3.3 pieces per minute, maintaining quality by adaptively controlling temperature and humidification duration. Among the good cups from this run, 78% exhibited a wrinkle count between 80 and 100, indicating good quality. The remaining 22% of the cups can be classified as very good quality with wrinkle counts over 100. The TCM was able to incrementally increase the die temperature up to 120°C, as required for each production cycle. Furthermore, by adjusting blankholder forces in response to moisture content fluctuations in the blanks, the TCM successfully formed quality-compliant cups. A second test series, comprising a total of 20 samples, was conducted to evaluate the performance of fixed parameter settings and the TCM under disturbed production conditions. To simulate these disturbances, a spray bottle was used to manipulate the material moisture content of selected blanks, resulting in significant changes to the material properties of subsequent blanks. The conventional fixed parameter setups failed to produce quality-compliant cups from the manipulated blanks. In contrast, the TCM demonstrated adaptability by modifying process parameters, particularly the blank holder force, to compensate for altered material properties. Consequently, both manipulated and undisturbed blanks were successfully formed into quality-compliant cups. This implies that the primary advantage of using TCM lies in its ability to adapt to unexpected events, such as unusually high material moisture content. In this scenario, the TCM system adjusted the process parameters in real-time, enabling the production of cups to meet quality requirements. In contrast, conventional control systems resulted in rejected products due to major defects. This adjustment enhanced process robustness against fluctuating environmental conditions.

The presented work successfully developed and deployed an intelligent process control system on a demonstrator that produced paperboard cups from a continuous web using compression-drawing. The comprehensive framework for its development combined digital twins and ML in a methodological approach based on four modules (see Chapter 4), resulting in Technology Control Model (TCM). The produced TCM not only successfully predicted paperboard cup quality from key process parameters but also optimized and adjusted them in coordination with the industrial control system. It was found that under stable initial conditions, the TCM showed no significant advantage over conventionally fixed parameters. However, TCM demonstrated its strength in the provocation tests in which the material moisture of individual blanks was specifically changed. A primary limitation found was that the TCM was trained solely on synthetic data where the moisture and temperature ranges were limited (<80°C and >6.7%). Despite this limitation, the TCM functioned as intended, predicting and optimizing parameters to ensure robust paperboard compression-drawing, while paving the way for innovative and sustainable manufacturing. Future improvements will focus on integrating real production data and product tracking.

Acknowledgements

The authors gratefully acknowledge financial support from the German Federal Ministry for Economic Affairs and Climate Action (BMWK) via grant number 01IF22573N (Project: commissioning of intelligent production systems).

References

[1] Abenghal, L., Bley, J., Tolnai, B., Njamen, G., Chabot, B., 2025. Development of a new sustainable packaging paper based on cellulose filaments and refined kraft pulp. Future Foods 11, 100540.

[2] 2017. Gesetz über das Inverkehrbringen, die Rücknahme und die hochwertige Verwertung von Verpackungen: Verpackungsgesetz: VerpackG.

[3] Marek Hauptmann, 2010. Die gezielte Prozessführung und Möglichkeiten zur Prozessüberwachung beim mehrdimensionalen Umformen von Karton durch Ziehen: Dissertation.

[4] Lenske, A., Müller, T., Penter, L., Schneider, M., Hauptmann, M., and Majschak, J., 2017. Evaluating the factors influencing the friction behavior of paperboard during the deep drawing process,. BioRes. 12(4), 8340-8358.

[5] Müller, T., Lenske, A., Hauptmann, M., and Majschak, J. P., 2017. Analysis of dominant process parameters in deep-drawing of paperboard. BioRes. 12(2), 3530-3545.

[6] Alava, M., Niskanen, K., 2006. The physics of paper. Rep. Prog. Phys. 69 (3), 669–723.

[7] Dipl.-Ing. Dominik Huttel, Dipl.-Ing. Paul-Philipp Post, 2015. Anwendung wirkmedienbasierter Verfahren_Dominik Huttel. Das IGF-Vorhaben 17788N der Forschungsvereinigung Kuratorium für Forschung und Technik der Zellstoff- und Papierindustrie.

[8] Alexey Vishtal, „Formability of paper and its improvement“. Dissertation, Tampere.

[9] Link, P., Penter, L., Rückert, U., Klingel, L., Verl, A., Ihlenfeldt, S., 2025. Real-time quality prediction and local adjustment of friction with digital twin in sheet metal forming. Robotics and Computer-Integrated Manufacturing 91, 102848.

[10] Digital Twins and Their applications_Sarow.

[11] Yao, J.-F., Yang, Y., Wang, X.-C., Zhang, X.-P., 2023. Systematic review of digital twin technology and applications. Visual computing for industry, biomedicine, and art 6 (1), 10.

[12] Tao F, Zhang M, Liu Y, Nee A. Digital twin driven prognostics and health management for complex equipment, in: , CIRP Anals, vol. 67.

[13] Jia, W., Wang, W., Zhang, Z., 2022. From simple digital twin to complex digital twin Part I: A novel modeling method for multi-scale and multi-scenario digital twin. Advanced Engineering Informatics 53, 101706.

[14] Attaran, M., Celik, B.G., 2023. Digital Twin: Benefits, use cases, challenges, and opportunities. Decision Analytics Journal 6, 100165.

[15] Zhang, J., Ellwein, C., Heithoff, M., Michael, J., Wortmann, A., 2025. Digital twin and the asset administration shell. Softw Syst Model 24 (3), 771–793.

[16] Lechler, T., Fischer, E., Metzner, M., Mayr, A., Franke, J., 2019. Virtual Commissioning – Scientific review and exploratory use cases in advanced production systems. Procedia CIRP 81, 1125–1130.

[17] Striffler, N., Voigt, T., 2023. Concepts and trends of virtual commissioning – A comprehensive review. Journal of Manufacturing Systems 71, 664–680.

[18] Oppelt, M., Urbas, L., 2014. Integrated virtual commissioning an essential activity in the automation engineering process: From virtual commissioning to simulation supported engineering, in: IECON 2014 - 40th Annual

Conference of the IEEE Industrial Electronics Society. IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society, Dallas, TX, USA. 29.10.2014 - 01.11.2014. IEEE, pp. 2564–2570.

[19] Jhaveri, R.H., Revathi, A., Ramana, K., Raut, R., Dhanaraj, R.K., 2022. A Review on Machine Learning Strategies for Real-World Engineering Applications. Mobile Information Systems 2022, 1–26.

[20] Chen, T., Sampath, V., May, M.C., Shan, S., Jorg, O.J., Aguilar Martín, J.J., Stamer, F., Fantoni, G., Tosello, G., Calaon, M., 2023. Machine Learning in Manufacturing towards Industry 4.0: From 'For Now' to 'Four-Know'. Applied Sciences 13 (3), 1903.

[21] Kulkarni, B., Tayde, S., Chapke, Y., Vyavahare, S., Badadhe, A., 2025. Machine learning techniques in monitoring and controlling friction stir welding process: a critical review. Discov Appl Sci 7 (6).

[22] Rojek, I., Kopowski, J., Lewandowski, J., Mikołajewski, D., 2024. Use of Machine Learning to Improve Additive Manufacturing Processes. Applied Sciences 14 (15), 6730.

[23] Nahvi, S., Schumann, M., Dix, M., Klimant, P., 2024. Intelligent parameter determination and quality control of an injection molding process. Procedia CIRP 126, 721–726.

[24] L. Berthold, G. Steinert, J. -P. Majschak. Compression-Drawing of Moisturized Paperboard: Experiments on Moisture and Temperature Dependencies. [accepted and in printing]. BioResources.

[25] Klara Liesegang, Patrick Link, Georg Steinert, Lena Berthold, Steffen Ihlenfeldt, 2026. Quality Assessment and Fast Geometry Prediction in Paperboard Forming. [under review]. The International Journal of Advanced Manufacturing Technology.

Biography

Muhammad Faisal Yaqoob (*1996) works as a research associate at the Fraunhofer Institute for Machine Tools and Forming Technology (IWU) since 2022. He completed his master's in computational science and engineering from Rostock University in 2022. His research focuses on digital manufacturing solutions, industry 4.0, digital twins and virtual commissioning.

Klara Liesegang (*1999) has been working as a research associate at the Fraunhofer Institute for Machine Tools and Forming Technology (IWU) since 2024. She completed her degree in applied mathematics at the TU Bergakademie Freiberg in 2023. She works on digital twins, process models, and simulations for compression-drawing processes.

Stephan Kronenberger has been working since 2024 as a technical staff member at the Fraunhofer Institute for Machine Tools and Forming Technology (IWU). He earned his Diploma (FH) in General Mechatronics from Wilhelm Büchner University in 2012. His work focuses on digital plant twins and virtual commissioning.

Dr.-Ing. Christer-Clifford Schenke is working at the Fraunhofer Institute for Machine Tools and Forming Technology (IWU) in Dresden since 2020. As a graduate in mechatronics, he completed his doctoral thesis on the expansion of system simulation in the field of forming technology and now heads the IIoT Control Systems and Technical Cybernetics department.

Prof. Dr.-Ing. Steffen Ihlenfeldt has been Director of the Fraunhofer Institute for Machine Tools and Forming Technology IWU and Head of the Chair of Machine Tools Development and Adaptive Controls at the Technical University Dresden. He focuses on cyber-physical production systems, machine data analysis, and human-robot collaboration, while being an Associate Member of CIRP, MHI, and WGP.

Lena Berthold (*1995) is Research Associate at the Institute of Natural Materials Technology at TUD Dresden University since 2019. The same year she completed her Diploma in Mechanical Engineering. Her research focusses on processes and machinery for sustainable packaging solutions.

8th Conference on Production Systems and Logistics

Smart Factory Framework for Real-Time Inventory Control and Process Optimization

Tom Wanyama
McMaster University, Hamilton, Canada

Abstract

This paper introduces a Smart Factory framework that integrates real-time inventory control with process optimization using Industrial Internet of Things (IIoT) technologies. The framework is designed to enhance operational efficiency by providing seamless visibility of material flow, proactive replenishment, and automated feedback to production systems. At its core is a weight-based inventory monitoring system that continuously measures component quantities and transmits the data through an IIoT network to the process control system. To complement this functionality, custom QR codes enable rapid parts identification, retrieval, and verification, eliminating human errors associated with manual entry and tracking. The system addresses persistent challenges in manufacturing environments, including unexpected downtime from stock-outs, inefficient material handling, and excess worker motion. By enabling real-time data acquisition and analytics, it allows both immediate process adjustments and retrospective analysis to optimize inventory policies and production schedules. The pilot implementation conducted at the SEPT Learning Factory at McMaster University demonstrated substantial performance improvements: elimination of data entry errors, significant reduction in workstation idle times, and a more than 10 percent increase in overall equipment efficiency (OEE) across all stations. Beyond immediate operational gains, the proposed framework exemplifies the broader transformation toward data-driven, adaptive manufacturing ecosystems. It demonstrates how tightly coupling inventory visibility with process control not only improves throughput and product quality but also enhances decision-making, transparency, and resource utilization across the supply chain. This work contributes to the ongoing evolution of smart manufacturing by providing a practical and scalable model for implementing real-time inventory intelligence as a foundation for sustainable, optimized production systems.

Keywords

Smart Factory, Industrial Internet of Things (IIoT), Real-Time Inventory Control, Process Optimization, Learning Factory, Operational Efficiency, overall equipment effectiveness (OEE), Smart Manufacturing.

1. Introduction

In the current business landscape, the strategic importance of effective supply chain management is continually gaining recognition. In fact, companies are increasingly directing their attention towards optimizing materials flow and supply chain processes to maximize profitability [4]. A key segment of the materials supply chain is the movement of components, raw materials, and other inventory through the processing cycle to create finished products and shipping of products to consumers. This constitutes the materials last mile. According to Boysen et al., 2021, in manufacturing, the last mile is the channel that connects the final (client) assembly equipment with the access node of the provider (carrier) warehouse [2].

DOI: https://doi.org/10.15488/20976
ISSN: 2701-6277

Therefore, the focus of this paper is the implementation and operation of systems that monitor and control materials inventory in manufacturing facilities.

Material flow is arguably one of the most important functions in a manufacturing because it is a major component of creating products from materials and shipping products customers. In fact, efficiency in material flow results in increased profits [4, 9]. Companies leverage potential efficiencies in last mile logistics through material flow analysis, resulting in capital savings and substantial reductions in storage, transportation, and handling costs. Beyond direct financial benefits, optimizing material flow minimizes energy consumption, waste generation, and floor space requirements, consequently reducing overall production time. In the long term, such streamlined material flow systems align with sustainable production practices, enabling compliance with forthcoming transparency and sustainability requirements [3]. This approach not only diminishes the company's carbon footprint but also ensures the judicious use of raw materials.

We have developed a weight-based parts inventory monitoring and control system, leveraging the capabilities of the Industrial Internet of Things (IIoT). This innovative system addresses various challenges, such as (1) eliminating downtimes caused by unforeseen stock-outs at workstations, (2) preventing human errors that result in the incorrect delivery of parts to workstations and subsequent downtimes, (3) improving product quality by ensuring that operator using the right materials in right qualities, and (4) minimizing motion waste by reducing unnecessary movements of workers searching for required parts at their workstations. The system incorporates custom QR codes to simplify inventory tracking and retrieval processes. Furthermore, the data generated by the system serves a dual purpose. Firstly, it allows for real-time optimization of parts flow within manufacturing facilities, enhancing overall operational efficiency. Secondly, it facilitates the retroactive optimization of stock levels, contributing to more informed decision-making regarding inventory management.

Our pilot implementation of the inventory monitoring and control system at SEPT Learning Factory has successfully eradicated data entry errors, optimized parts delivery, and minimized workstation downtimes. This has led to a remarkable increase of over 10% in the overall equipment efficiency across all workstations. Leveraging is IIoT features, we have seamlessly integrated its information into the process control system, enhancing the quality of our products. The remainder of this paper is organized as follows: Section 2 delves into the motivations driving the development of our inventory monitoring system, while Section 3 provides comprehensive details of the system, encompassing its evaluation and potential future enhancements. Section 4 concludes the paper.

2. Motivation

For a company to thrive, it is imperative that its material flow system operates seamlessly, ensuring maximum efficiency. However, the material management of many companies grapples with recurring challenges. The major persistent challenges of material flow in manufacturing facilities include:

1. Warehouse Capacity Issues: The layout of a warehouse holds paramount importance in effectively utilizing space. For instance, the strategic placement of low-running inventory at the front and high-running inventory at the back may lead to operational inefficiencies, demanding more effort and time to access items.
2. Inventory Inaccuracy: Discrepancies in inventory tracking often stem from errors in recording incoming or outgoing goods, resulting in inaccurate stock levels and unplanned down times.
3. Unidentified Locations of Inventory: The arbitrary storage of goods can complicate the location of items during shipment, causing delays and operational inefficiencies.

In response to these challenges, companies are increasingly turning to advanced technologies to streamline and enhance their material management processes.

2.1 Material Monitoring Systems

As part of this research, we set out to carry out a study aiming at revieing and comparing various inventory monitoring systems to assess their effectiveness in a manufacturing environment. This involved evaluating how well these systems performed in terms of tracking inventory, reducing errors, and optimizing inventory management processes. By establishing this objective, the study aimed to identify the most suitable systems that could enhance operational efficiency and productivity in manufacturing settings. Next, it was essential to establish selection criteria for choosing the inventory monitoring systems to be reviewed. The criteria included several key aspects: technological features, such as the presence of advanced technology like IoT, real-time monitoring, and automation; functionality, including a range of capabilities like inventory tracking, error reduction, and optimization; industry relevance, ensuring suitability for use in manufacturing and similar industries; market presence, focusing on well-known or widely used systems; and innovation, highlighting systems that offered unique or innovative solutions. Preliminary research was then conducted to identify potential candidates, involving a literature review, consultation with industry experts, and online searches for product information. Based on this research, a long list of candidates was compiled, which was then refined by applying specific inclusion and exclusion criteria, such as relevance to manufacturing, documented case studies, and availability of technical information, while excluding systems with limited functionality or outdated technology.

The next steps involved a detailed evaluation of the shortlisted systems based on the defined criteria. This included a thorough analysis of their features and functionalities, assessment of usability and integration capabilities, cost-benefit analysis, and review of performance metrics like accuracy and reliability. From this detailed evaluation, the most promising systems were selected for an in-depth review. In this study, the systems selected were SmartShelf from Mettler Toledo, StockVUE, Limble computerized maintenance management system, and SmartBin IoT material monitoring solution from Bassard. They key features of theses inventory monitoring systems are summarized as follows:

- SmartShelf from Mettler Toledo [7], for instance, offers a comprehensive parts management solution designed to tackle issues related to parts tracking and usage. Notable features of SmartShelf system include instant notifications for low part quantities, real-time reporting, and automatic parts usage tracking. The system has scales that connect to PCs through PCB Speedy boards (Scale Boards) connected to a converter box using Ethernet cables. This enables the materials to be monitored using a proprietary software application that communicates with the scales using a proprietary protocol. The data can be sent to a cloud server to enable local and remote monitoring of materials.
- Another inventory monitoring system, StockVUE, leverages cloud-connected smart scales and IoT technology to revolutionize inventory management as shown in Figure 1. The system aims to minimize the need for travel to remote sites, eliminate stock-outs, and enhance overall operational efficiency [8]. The StockVUE web application provides remote access to inventory in real time. The stand-alone web application can be accessed by employees, customers, and partners to keep the entire value chain in sync. Easy to use APIs provide connectivity between StockVUE Cloud and ERPs or warehouse management systems making it compatible with existing systems (Figure 1).
- Limble is a cloud-based computerized maintenance management system (CMMS) that aids in the management of maintenance work by automating work scheduling, monitoring work history, generating reports, and organizing assets, including materials [6]. It provides a comprehensive solution for parts management, addressing issues related to parts tracking and usage. Notable features of Limble's system include instant notifications for low part quantities, real-time reporting, automatic tracking of parts usage, and integration of barcode lookup for efficient inventory tracking. It is important to note that real-time weight-based material monitoring is not a native feature of Limble, but it can be integrated with in-house systems to provide this functionality.
- SmartBin IoT from Bossard uses wireless, battery-powered weighing scales that transmit inventory data to a cloud-based platform. The system supports threshold-based replenishment, automated

transmission of inventory data to a centralized application, and consolidation of replenishment orders into scheduled deliveries. While effective for warehouse-level monitoring, the architecture is primarily cloud-centric and does not provide deterministic, low-latency access suitable for real-time process control integration [1].

Although weight-based inventory monitoring has received limited direct coverage in the academic literature, existing work on IIoT-enabled material visibility, edge-based sensing, and cyber-physical production systems provides a strong conceptual foundation for the proposed approach.

2.2 Strengths and Weaknesses of Inventory Monitoring Systems

The optimal configuration for inventory monitoring systems in manufacturing allows for adaptable processing that can be modified to accommodate variations in product volume, quantity, and type. The system should interact with human input and other data and processes, ensuring predictability and seamless resolution of potential issues. A well-crafted inventory monitoring system empowers manufacturers to reduce waste and improve production capabilities; and developing such a system must be guided by the following three interconnected principles:

1. Planning: The engineered material flow requires a precise and intentional outline of material movement, including accurate data on components and materials, and identification of each step in the manufacturing process.
2. Communication: Smooth operation relies on effective communication and signals, which can take various forms such as floor arrows, overhead lighting, barcodes, colored boxes, and numbered signs.
3. Tools: Tool selection depends on the manufacturing environment. Processes should be analyzed first, and equipment should then be chosen to align with the specific needs of the operation.

All the reviewed material inventory monitoring systems are flexible enough to cater to the specific plans of manufacturing facilities. However, only Limble CMMS is not associated with a specialized material update tool (weighing scales). Surprisingly, only SmartShelf from Mettler Toledo provides local access to the data. But unfortunately, this access is facilitated through a proprietary software application that retrieves data from the weighing scales using a proprietary communication protocol, making it difficult to integrate SmartShelf with local process control systems.

Data access for all reviewed inventory monitoring systems to create Industrial Internet of Things is through cloud servers. While this is advantageous for remote data access and accessibility for powerful computing resources, it introduces the following limitations:

- Long Latency: Cloud-based systems move processing further away from data sources, increasing the distance that data needs to travel. This heightened distance results in increased latency and degraded response times, crucial for using the data in process control, such as ensuring product quality by real-time verification that the parts used meet customer specifications.
- Privacy and Security: Local data storage and processing allow sensitive data to be processed locally, minimizing the need to transmit sensitive information to the cloud. This approach enhances privacy and security.
- Offline Operation: Eliminating the need for cloud computing enables devices to perform computations locally even when they are not connected to the internet. This is particularly useful for industrial automation systems requiring data on the availability of materials.
- Scalability: Local data storage and processing can distribute computational tasks across multiple edge devices, leading to improved scalability. This is particularly advantageous in scenarios where many devices need to work together without overwhelming a centralized cloud infrastructure.
- Cost Savings: Local data storage and processing can reduce the costs associated with transmitting large volumes of data to the cloud, especially in scenarios where data storage and processing are performed locally. This can result in cost savings related to bandwidth usage and cloud service fees.

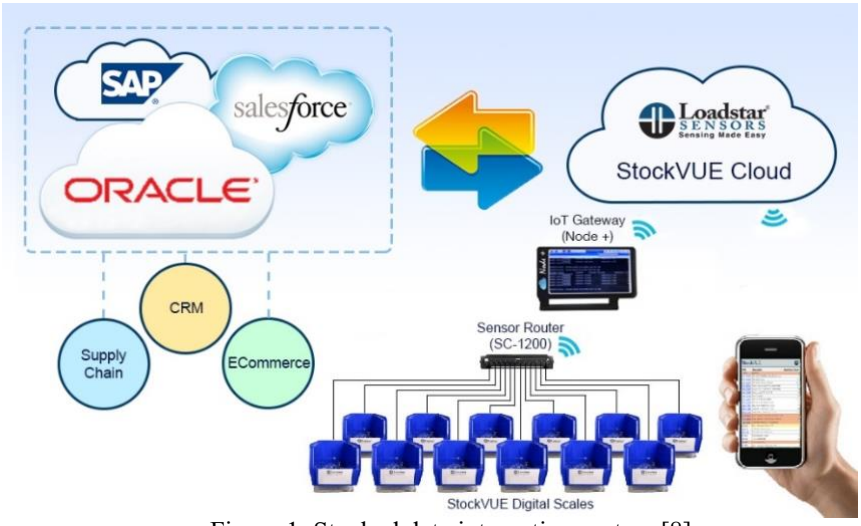


Figure 1: Stocked data integration system [8]

3. EdgeSensor: Smart Material Monitoring System

At SEPT Learning Factory, we have developed a cloud and edge computing-based inventory monitoring system that utilizes weight scales to monitor and update inventory at various stages of production. It also allows for the integration of other tools, such as barcode scanners, for inventory updates. Figure 2 shows that the scales use a standard industrial automation network protocol, Modbus TCP, for communication with the process controllers (PLCs) and local databases, and MQTT for data transmission to public and proprietary cloud servers. It's crucial to note that edge computing and cloud computing are not mutually exclusive; and for our system they can complement each other in a hybrid architecture where data processing and sharing for products assembly and quality control is done at the edge, and long-term data storage and remote access tasks are offloaded to the cloud. Our system is designed not only to support materials monitoring in warehouses but also to support real-time monitoring of material consumption and decision making at workstations to support quality improvement in Industry 4.0 enabled manufacturing environment. For this reason, the ability to perform computations locally even when the system is not connected to the internet, enabled by edge computing, is of paramount importance. Therefore, we call our system EdgeSense: Smart Material Monitoring System.

Weight-based inventory monitoring was selected due to its robustness, simplicity, and suitability for high-mix manufacturing environments. Unlike vision-based or count-based systems, weight-based sensing is insensitive to lighting conditions, part orientation, and operator handling variability. It enables continuous, non-intrusive measurement without requiring manual interaction or precise part alignment. While the approach is sensitive to material density variation and requires calibration, these limitations are mitigated through bin-specific configuration and filtering at the edge layer. The method therefore provides a pragmatic balance between accuracy, cost, and integration complexity, making it well suited for real-time manufacturing environments.

The smartness of EdgeSense is three folds: Firstly, it allows the operator at a workstation to arrange the materials in any configuration because the association between material bins and material stations in not hard coded in the system. Instead, the operator dynamically attached material bins to material stations. Secondly EdgeSense updates the list of materials used at workstations based on the Manufacturing Execution System (MES) product routing. That is, materials consumed at workstations are not hard coded in the system. Thirdly, EdgeSense places order to the warehouse or materials supermarket for workstations automatically. If warehouse inventories reach predefined levels, the system places orders to suppliers automatically, supporting supply chain management as shown in Figure 3. EdgeSense sends data to cloud servers to enable remote data access and integration of legacy MES and Enterprise Resource Planning (ERP) systems such as

SAP and salesforce (Figure 2). The data stored in cloud servers provide opportunities for retractive studies of stock levels for informed decision-making and inventory optimization and management.

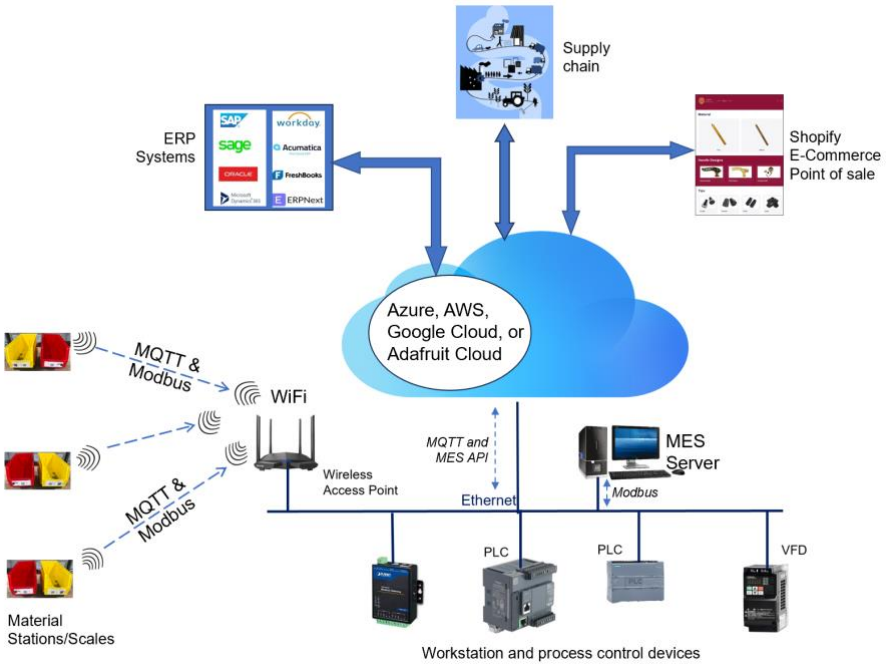


Figure 2: Network architecture of EdgeSense

3.1 Evaluation of EdgeSense

We have deployed EdgeSense in SEPT Learning Factory located in W. Booth School of Engineering Practice and Technology, Faculty of Engineering - McMaster University, to monitor warehouse and workstation inventory. Figure 3 shows the distributed intelligence of EdgeSense, with the scales of the material stations collecting the weight of bins that they send to the workstation PLCs. The PLCs obtain the list of materials to be consumed at the workstations from the MES, which in turn receives the product specification from the orders placed through the points of e-commerce sales on Spotify.

After the operator attaches the material bin to the materials station by scanning the QR-code of the station followed by the QR-code of the bin, the inventory in the bin is displayed on the operator Human Machine Interface (HMI) and the location of materials is updated in the MES that manages materials in the manufacturing facility. During product assembly, the press and glueing machines controlled by the workstation PLC shown in the example workstation in Figure 3, are activated only when EdgeSense indicates that the right part has been picked from the materials bin.

In addition, the amount of glue discharged is automatically sent by the part being glued as indicated by EdgeSense. This is an example of how EdgeSense ensures product quality. In general, the automated flow of information among the point of sale, MES, EdgeSense material bins and stations, and the process controls is the backbone of Industry 4.0.

Our initial trial of EdgeSense at the SEPT Learning Factory has yielded impressive results. By deploying EdgeSense, we have successfully eliminated errors associated with manual data entry, optimized the delivery of parts, and minimized downtimes at workstations. As a result, we've seen a remarkable improvement of over 10% in the overall equipment efficiency across all our workstations. Overall Equipment Effectiveness (OEE) was evaluated using standard availability, performance, and quality metrics derived from MES event logs and PLC cycle data. Availability improvements were primarily attributed to the elimination of unplanned workstation stoppages caused by material shortages. Performance gains resulted from reduced operator walking time and more stable cycle execution, while quality improvements were observed through the prevention of incorrect part usage enabled by QR-code verification. Although the evaluation was

conducted within a Learning Factory environment, the measurement methodology aligns with established industrial OEE practices and is transferable to production settings with similar material-flow characteristics. This OEE increase can be attributed to the reduction in wait times caused by shortages of parts both at workstations and in the warehouse. Furthermore, leveraging the advanced IIoT capabilities of EdgeSense has allowed us to seamlessly integrate inventory information into our process control system. This integration has significantly enhanced the quality of our products by ensuring that the right parts are available at the right time and place. Overall, the implementation of EdgeSense has not only improved efficiency but has also contributed to the overall quality and reliability of our manufacturing process at the SEPT Learning Factory.

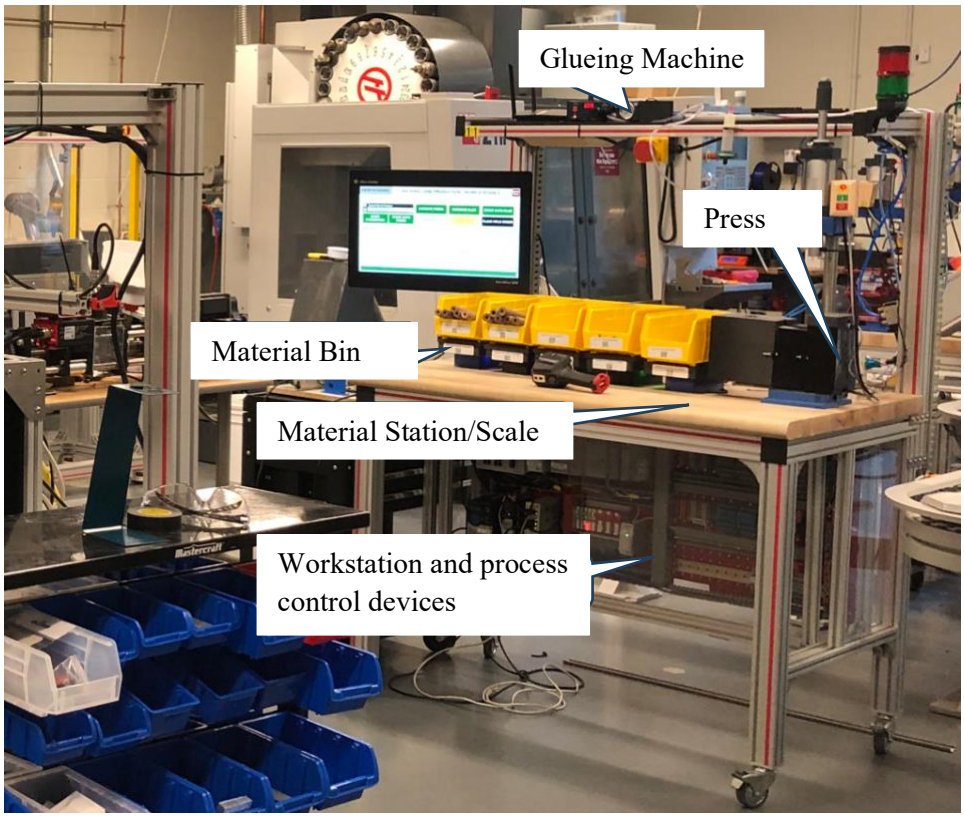


Figure 3: SEPT Learning Factory: EdgeSense Workstation Example

3.2 Future of EdgeSense

Having successfully concluded the pilot study, our current focus is on extending the deployment of EdgeSense to diverse manufacturing environments, aiming to evaluate its adaptability and performance under different operational conditions. Concurrently, we are exploring alternative network architectures, specifically investigating the feasibility and effectiveness of both wired and wireless Modbus RTU communication protocols between materials stations and materials Programmable Logic Controllers (PLCs). In addition to this, we are pursuing the development of an innovative approach where EdgeSense would be integrated with Manufacturing Execution Systems that reside on cloud servers. This initiative is poised to improve the adaptability of EdgeSense, increasing its flexibility, scalability, and accessibility while ensuring robust data management and security measures by minimizing the need to send sensitive information to the cloud.

4. Conclusion

The implementation of EdgeSense, a weight-based parts inventory monitoring and control system represents a significant advancement in manufacturing operations. The system harnesses the capabilities of the Industrial Internet of Things (IIoT), to effectively addresses critical manufacturing challenges such as stock-

outs, human errors, and motion waste, leading to enhanced operational efficiencies. The integration of custom QR codes streamlines inventory tracking and retrieval processes, while the generated data enables real-time optimization of parts flow and retroactive analysis for informed decision-making in inventory management.

The pilot implementation at SEPT Learning Factory yielded tangible benefits, including the elimination of data entry errors, optimized parts delivery, and minimized workstation downtimes, resulting in a notable increase in overall equipment efficiency. Moreover, the seamless integration of information into the process control system through IIoT features contributes to enhancing product quality. This successful deployment underscores the importance of effective parts inventory tracking in manufacturing, emphasizing transparency, improved inventory control, and overall profitability. Furthermore, it aligns with the industry's evolving focus on optimizing supply chains and maintaining well-managed warehouses to ensure maximum efficiency.

References

[1] Bassard, Smartbin and SmartCloud Solution, 2024. available as of February 21st at https://www.bossard.com/ca-en/smart-factory-logistics/systems/smartbin-cloud/.

[2] Boysen, N., Fedtke, S., Schwerdfeger, S. 2021. Last-mile delivery concepts: a survey from an operational research perspective. OR Spectrum (43) 1.

[3] FasterCapital, 2024. Energy Efficiency: Sustainable Production Practices for Reduced Costs, Available as of February 21st, 2024 at: https://fastercapital.com/content/Energy-Efficiency--Sustainable-Production-Practices-for-Reduced-Costs.html.

[4] Guzenko, A.V., Guzenko, N.V., 2022. Process optimization for last mile logistics. Transp. Res. Procedia, (63), pp. 1700–1707.

[5] Cullen, J. M., Cooper, D. R., 2022. Material Flows and Efficiency, Annual Review of Materials Research, (52), pp 525-559.

[6] Limble, 2024. cloud based computerized maintenance management system, available as of February 21st at https://limblecmms.com/.

[7] Mettler Toledo, 2024. SmartShelf - Inventory Management System -Shelf scales enabling automated, real-time stock status and replenishment, available as of February 21st at: https://www.mt.com/ca/en/home/products/Industrial_Weighing_Solutions/bench-scales/smart-shelf.html.

[8] Stockvue, 2024. Inventory that Counts Itself - Track Your Inventory Remotely, Autonomously, in Real-Time https://www.stockvue.net/.

[9] Thanou, E., Matopoulos, A., 2021. Improving efficiency of material flows in an automotive assembly plant: A case study, Procedia, (35), pp. 959-967.

Biography

Dr. Tom Wanyama is the Director of the SEPT Learning Factory and Chair for the Bachelor of Technology in Automation Systems Engineering Technology program, Faculty of Engineering at McMaster University. With over 20 years of university teaching experience, Dr. Wanyama's expertise spans from electrical engineering to automation systems integration. His research focuses on manufacturing, IoT/AI and decision-making, and computer-based teaching tools.

8th Conference on Production Systems and Logistics

Decoupling Mode And Effects Analysis: Enhancing Manufacturing Resilience Through Systematic Risk Identification

Florian Pleier[1], Simon Rekers[2], Andreas Röder[3], Daniel Beck[4], Ben Lüskow[5], Kirsten Tracht[5], Johannes Schilp[1,4]

[1]*University of Augsburg, Institute of Applied Computer Science, Department of Digital Manufacturing, Augsburg, Germany*
[2]*DX FACTURE, Aachen, Germany*
[3]*IPRI – International Performance Research Institute, Stuttgart, Germany*
[4]*Fraunhofer IGCV – Fraunhofer Institute for Casting, Composite and Processing Technology, Augsburg, Germany*
[5]*University of Bremen, Bremen Institute for Mechanical Engineering (bime), Bremen, Germany*

Abstract

Global production systems are increasingly vulnerable due to climate change, pandemics, geopolitical tensions, and trade restrictions. Moreover, there is an observable long-term decoupling of the global economy to reduce interdependencies. These factors have the potential to disrupt established supply networks, reduce operational flexibility, and challenge the long-term competitiveness of manufacturing companies. To address these vulnerabilities, manufacturing companies are required to establish resilient production systems. However, existing methods often fail to systematically assess the specific risks associated with decoupling. This article proposes a systematic approach to identify potential vulnerabilities in production systems. The Decoupling Mode and Effects Analysis (DMEA) aims to enhance organizational resilience by identifying and analyzing potential risks associated with decoupling. This approach adapts the established Failure Mode and Effects Analysis (FMEA) framework to this new context, integrating a structured review of manufacturing processes, relevant decoupling scenarios, and their potential effects on system performance. By prioritizing risks according to severity, likelihood, occurrence velocity, and detectability, the DMEA provides a practical tool for enhancing organizational resilience in the face of global economic fragmentation. The DMEA enables the early identification and quantification of vulnerabilities related to decoupling, allowing companies to implement targeted resilience measures proactively. Beyond the boundaries of individual companies, the DMEA can also be extended to systematically assess risks within interconnected value creation networks, thereby supporting cross-company resilience assessments.

Keywords

Decoupling; DMEA; FMEA; Resilience; Manufacturing; Production System; Risk Assessment

1. Introduction

Global production systems are increasingly exposed to vulnerabilities arising from geopolitical tensions, trade restrictions, climate risks, and shifting economic policies. These developments threaten the stability of supply networks and increase the structural uncertainty under which manufacturing companies operate [1]. In this context, decoupling refers to the deliberate reduction of interdependencies between organizations to avoid over-reliance on specific regions or trading partners. Decoupling is driven by geopolitical tensions and economic factors beyond the control of affected companies [2,3]. Recent evidence shows that geopolitical initiatives toward economic self-sufficiency can have immediate operational impacts. For example, the strategic focus of some nations on strengthening domestic consumption has resulted in supply delays for more than half of the foreign companies operating within the country [4]. The consequences for

DOI: https://doi.org/10.15488/20977
ISSN: 2701-6277

manufacturing companies extend well beyond disrupted supply chains, affecting production planning, investment decisions, and workforce stability [5].

As a result, many companies are pushed into reactive crisis management, prioritizing short-term problem-solving over proactive and structured risk assessment [6]. Existing risk-management frameworks tend to focus on operational incidents and localized disruptions. However, these frameworks often fail to capture the structural vulnerabilities introduced by geopolitical decoupling, such as altered network dependencies, the loss of critical resources, or the increased exposure to unilateral political shifts [7]. This gap leaves manufacturing companies ill-prepared to anticipate and mitigate the strategic risks associated with an increasingly fragmented global economy.

To address this challenge, we propose a systematic method for identifying and evaluating decoupling-related risks in production systems. Our approach adapts the well-established Failure Mode and Effects Analysis (FMEA) framework by integrating a structured review of manufacturing processes, relevant decoupling scenarios, and their potential effects on system performance. By prioritizing risks according to their severity, likelihood, occurrence velocity, and detectability, the method enables companies to identify potential threats early. Moreover, it provides a practical tool for strengthening organizational resilience in periods of global economic fragmentation as well as for developing robust resilience strategies.

2. Theoretical Background (FMEA)

The development of the DMEA is conceptually based on the principles of the FMEA. The following section will thus provide the theoretical foundations of the FMEA, a well-established method in preventive risk management. This method is widely employed for the systematic analysis of complex systems and processes [8,9]. Its primary objective is to identify potential failure modes, their underlying causes, and their potential consequences at an early stage to derive appropriate measures to prevent failure modes. In this context, risk is defined as a combination of the occurrence probability and the severity of its consequences [10].

In quality and safety management, the FMEA is applied as a preventive measure to enhance the technical reliability of products and processes. It supports the systematic evaluation of potential failure modes and their effects and enables the comparison of different system or process variants with regard to their respective risk profiles. Key advantages of the method include its clear and logically structured procedure, its relatively low methodological complexity, and its broad applicability in nearly all technical domains [11].

The FMEA is carried out by an interdisciplinary team with clearly defined roles and responsibilities, which documents the analysis systematically using a standardized form. The initial step consists of a structural and functional analysis, in which the system or process under consideration is decomposed into its elements, interfaces, and functional interdependencies. Information regarding the structure and functionality of the system is typically derived from documents such as functional descriptions and process-flow diagrams. A failure analysis is then conducted, in which potential failure modes are systematically identified for each element within the scope of consideration. Subsequently, the potential effects of each failure mode are then described and linked to their respective causes, resulting in a comprehensible structure of failure modes, effects, and causes [12]. Risk assessment is performed using the Risk Priority Number (RPN), which is determined by the severity of the potential effects, the probability of occurrence, and the likelihood of detecting the failure before it manifests. The RPN serves as the basis for prioritizing risks and identifying where preventive or reactive measures are required. After the implementation of these measures, a reassessment is carried out to verify the effectiveness of the improvements and to determine the remaining residual risk [13].

The traditional focus of the FMEA on technical systems, products, or processes limits its applicability to the analysis of potential failure modes, their causes, and effects. Consequently, its suitability for examining organizational resilience or decoupling-related risks is restricted. The conventional FMEA assessment

dimensions are not designed to capture the systematic characteristics that shape organizational resilience, such as interdependencies, coupling structures, and adaptive capabilities [14].

3. State of the Art Methods

The state-of-the-art methods for assessing and prioritizing risks in production and supply networks can be condensed into three methodological lines. Firstly, there are classical hazard analyses, which focus on plant or process level. Secondly, there are resilience-oriented network approaches. Thirdly, there are frameworks for prioritization. As the FMEA has already been addressed in Section 2, it is only distinguished here.

Classical hazard methods - Fault Tree Analysis (FTA), Hazard and Operability Studies (HAZOP), BowTie analysis, and system-theoretic approaches (STPA) - primarily address technical and process hazards, cause-effect chains, and barriers at plant or parameter level [15–18]. FTA models top-down failure logic for structured root-cause analysis [16]; HAZOP systematically identifies deviations from target parameters and their consequences [17]; BowTie makes preventive and mitigative barriers explicit between causes and consequences [15]; and STPA extends the view to control structures and system constraints beyond components [18]. While these methods offer depth for operational and safety risks, they lack explicit concepts for network-wide occurrence velocity of disruptions and a standardized catalog of decoupling scenarios.

Resilience-oriented work in supply chain research partially addresses this gap by introducing time- and network-based metrics for vulnerability diagnosis. Time-based metrics such as Time-to-Recover (TTR) and Time-to-Survive (TTS) are practical, data-light analytical instruments that can be derived from ERP-proximate information such as inventory coverage, changeover or procurement times, transport or customs times [6,19]. TTR is the time it takes for a defined, acceptable performance level to be restored after a disruption occurs. TTS is the maximum amount of time that a system can remain functional after a disruption occurs, without any external intervention or measures, before reaching a critical threshold [14]. They quickly indicate whether a failure or decoupling scenario is operably bridgeable ($TTS \geq TTR$) and provide a direct operationalization of propagation dynamics: the smaller the TTS buffer relative to TTR, the higher the occurrence velocity of a shock along the production flow [19]. For manufacturing small and medium-sized enterprises (SMEs), these metrics are particularly suitable because they lead directly to actionable measures - inventory and capacity buffers, multisourcing, shorter changeover and qualification times - without requiring extensive data. [6,19]

Literature suggests that disruptions propagate along topologies and can create secondary bottlenecks. Concepts such as the “Ripple Effect” and “Viability” explicitly address these systemic dynamics and complement time-based metrics with structural insights [1,20]. In addition to simulation studies (discrete-event, agent-based), lightweight graph metrics - such as betweenness, minimum cuts, and path lengths - are used to identify structural criticality. Review papers synthesize quantitative resilience assessment methods and recommend phased adoption depending on data maturity [21].

In the domain of operations research, a comprehensive set of robust and stochastic optimization models facilitates the design of mitigation policies in the face of disruptions. Established frameworks provide concrete policies once vulnerabilities have been identified [22]. In the context of global decoupling, it can be concluded that, subsequent to the identification and prioritization of decoupling-related risks (including occurrence velocity), design and inventory decisions can be systematically optimized. At the same time, data and modelling requirements are comparatively high; consequently, operations research models are most effective as a follow-up deep dive once TTR/TTS and network heuristics have established the analytical focus.

For consistent prioritization under uncertainty, literature frequently highlights the benefits of multicriteria decision-making (MCDM) methods (e.g., AHP, TOPSIS, DEMATEL), fuzzy logic, and Bayesian methods.

They integrate qualitative and quantitative scales (severity, likelihood, detectability) to derive transparent ranking, similar to the FMEA [23]. However, FMEA-specific research also highlights several weaknesses of the classical RPN, such as scale inconsistencies and independence assumptions, which has led to numerous extensions, including occurrence velocity, which is especially important in the context of decoupling [24].

Established standards such as ISO/IEC 31010 and ISO 22301 introduce a comprehensive toolbox of risk-assessment techniques (e.g., BowTie, FTA, Monte Carlo analysis) and embed business-continuity management into a structured framework of scenario planning, roles, exercises, and performance indicators [25,26]. Both standards are compatible with resilience assessments but do not offer decoupling-specific scenario catalogs or mechanisms for capturing the network-wide occurrence velocity of a risk. Therefore, geopolitical risk indices can be employed to indicate country-specific tensions over time. One example is the Geopolitical Risk Index (GPR), which is a text-based index that measures geopolitical risk at a country level and over time, based on the frequency of geopolitically relevant terms in major international newspapers [27]. However, their systematic integration into resilience assessment methods remain limited.

The central research gap relevant to this contribution lies in integrating a decoupling-specific scenario catalog, the explicit assessment dimension of occurrence velocity, and a process-step-level linkage to production systems. The DMEA addresses this gap by combining established risk logics (FMEA, see Section 2) with TTR/TTS-based stress tests, lightweight network indicators, and consistent prioritization (severity, likelihood, occurrence velocity, and detectability) focused on early identification and targeted reduction of vulnerabilities in manufacturing SMEs.

4. DMEA Method

The Decoupling Mode and Effects Analysis (DMEA) is a novel, structured analytical method designed to identify and evaluate decoupling-related risks within production systems and their corresponding networks. Its core objective is to enhance organizational awareness of such risks and to support manufacturing companies in deriving appropriate preventive actions. Conceptually, the DMEA builds upon the methodological principles of the FMEA but extends the analytical scope from technical failure modes toward business processes and their exposure to external political, economic, ecological, social, or technological developments that may trigger decoupling-induced disruptions.

The DMEA enables the systematic identification of business-relevant risks arising from critical dependencies, coupling structures, or insufficient adaptive capacity, and assesses their potential impact on organizational resilience. As such, it represents a methodological extension of the conventional FMEA approach, enriching it with resilience- and decoupling-related risk dimensions and enabling a structured analysis of these specific vulnerabilities. In addition, the DMEA is designed to maintain transparency, ensuring that risk assessments remain traceable across different organizational contexts.

As with FMEA, it is necessary to define some roles in advance of commencing work. The definition of clear roles forms the foundation of a consistent and traceable DMEA. A minimum set of core roles ensures that the analysis is methodologically correct, and that relevant expertise is integrated. Table 1 presents the essential roles and their responsibilities. Additional participants may be added depending on the complexity of the business processes or the depth of analysis required.

Table 1: DMEA roles and responsibilities

Role	Responsibility
DMEA Leader	Planning, moderation, and adherence to the methodology.
Process Owners	Provide detailed knowledge about the business processes considered in the system analysis and are involved in the conduction of the failure analysis.
Team Members	Cross-functional participants contributing technical and operational expertise to the failure analysis.
Recorder	Systematic documentation and subsequent, transparent communication of the DMEA results.

After defining the roles and responsibilities, the DMEA method consists of three successive steps: the system analysis, the failure analysis, and the quantitative assessment. Figure 1 provides an overview of the DMEA workflow, while the following sections describe each step in detail.

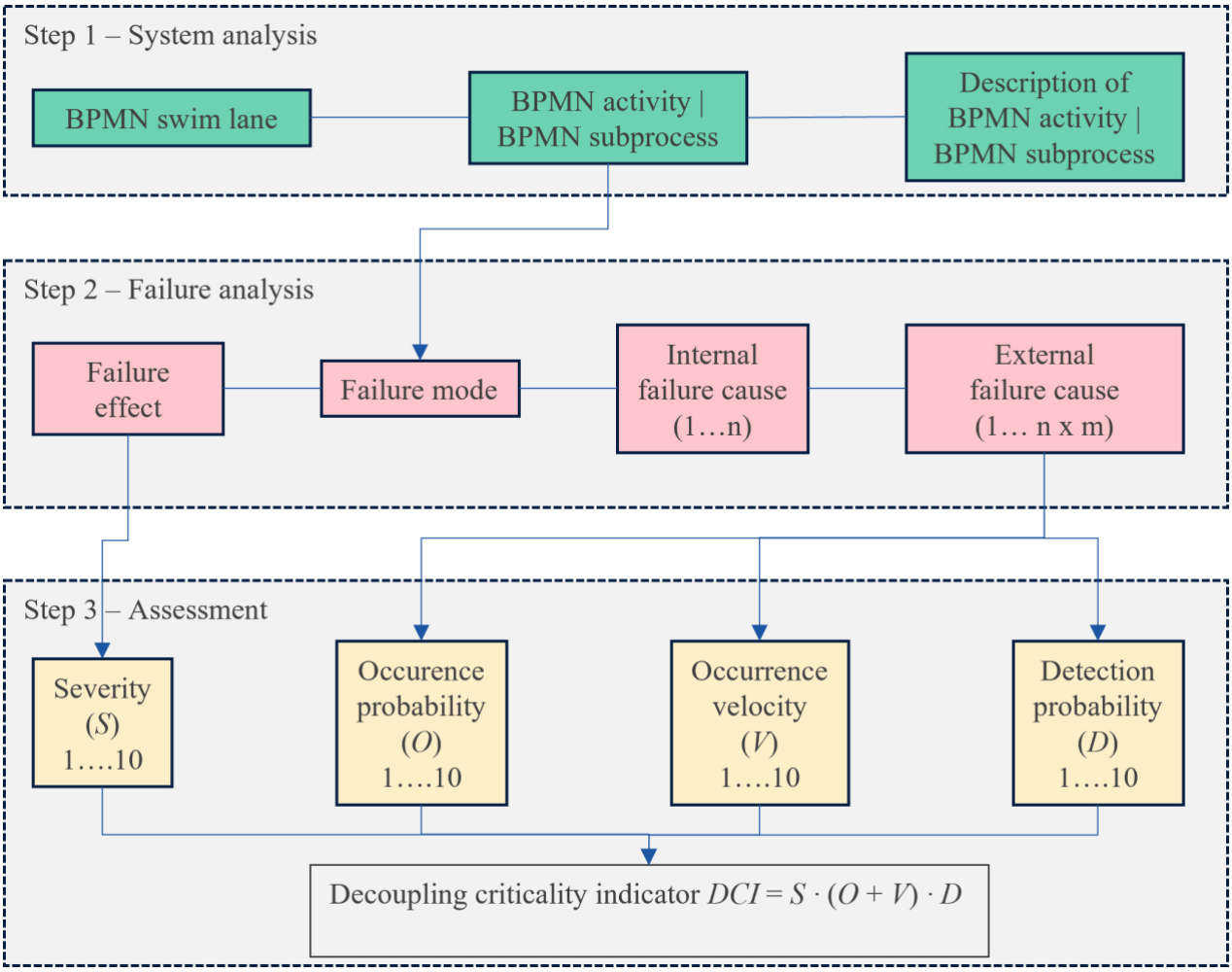


Figure 1: DMEA workflow

4.1 System analysis

The first step of the DMEA is the system analysis, which is illustrated by green elements in Figure 1. The system analysis forms the analytical foundation of the DMEA. Its purpose is to create transparency regarding the internal value-creation processes that may be affected by decoupling-related risks. To ensure a

standardized and comprehensible representation of business processes, the method uses Business Process Model and Notation (BPMN) [28].

For each entity involved, a dedicated BPMN swimlane is created, followed by the mapping of all relevant process steps and subprocesses. The resulting BPMN diagrams are then reviewed by the entire DMEA team. This ensures a common understanding of functional sequences, interfaces, and potential points of vulnerability within the process chain. The resulting validated BPMN model not only serves as the reference structure for the subsequent failure analysis but also creates a shared understanding among participants, which is essential for consistent risk identification.

4.2 Failure analysis

The second step of the DMEA is the failure analysis, which comprises the elements marked in red in Figure 1. The failure analysis builds directly on the BPMN-based system representation and is conducted in two consecutive parts.

In the first part, all BPMN activities and subprocesses are examined for potential failure modes. A failure mode in the DMEA context represents a deviation in the business process from the regular state that can be causally linked to an external failure cause, i.e., a trigger originating from outside the company (e.g., regulatory change, trade restriction, geopolitical escalation, environmental shock).

The second part describes the full causal chain, which is represented by the red elements in Figure 1. This chain is read from right to left, starting with the external failure cause, over an internal failure cause, to the failure mode, and finally the failure effect. It is important to mention that causal relationships may involve multiplicities. Each failure mode may have n internal failure causes. Each internal failure cause may be triggered by m external failure causes. Hence, up to $n * m$ external failure causes may exist for each failure mode. Failure effects must be expressed in monetary terms (e.g., financial loss, downtime cost, margin reduction) in order to quantify the severity. This requirement directly links the failure analysis to the subsequent severity rating.

To ensure transparency and traceability, the failure analysis is typically documented using an FMEA-like spreadsheet. Process owners and team members jointly develop the causal chain, while the recorder ensures the proper documentation, and the DMEA leader ensures adherence to the methodology.

4.3 Assessment

The third step of the DMEA is the assessment, which refers to the yellow elements in Figure 1. The aim of the assessment is to quantify the criticality of each decoupling-related risk. For this purpose, the DMEA introduces the Decoupling Criticality Indicator (DCI), which reflects the urgency of preventive action. The higher this indicator, the greater the need for proactive measures to mitigate the decoupling-related risk. To calculate the DCI, four individual quantitative assessments must be conducted. Each assessment is rated on a scale ranging from 1 (low risk) to 10 (high risk), corresponding to the range of the scale in the FMEA. The external failure cause is evaluated by using three different indicators. The occurrence probability O represents the likelihood or frequency of a specific external failure cause. The occurrence velocity V rates the time span between the occurrence of an external failure cause and the corresponding (internal) failure effect. The detection probability D describes how likely it is that an external failure cause will be detected before it affects the corresponding business process. Finally, a fourth indicator is applied to the failure effect. The severity S quantifies the expected monetary impact caused by a decoupling-related risk related to the external failure mode.

These four indicators are the basis for the calculation of the DCI:

$$DCI = S * (O + V) * D \quad (1)$$

In contrast to the FMEA, which multiplies the three factors of severity, occurrence probability, and detection probability to form the RPN, the DMEA adds the factor of occurrence velocity to the occurrence probability and multiplying their sum by the factors of severity and detection probability to calculate the DCI. This mitigates sensitivity issues typically associated with multiplying all four factors together, where minor changes in individual parameters can lead to disproportionately large shifts in the overall indicator. This was proven to be disadvantageous when comparing the overall risk indicators with each other and when analyzing the benefits of actions.

It should be noted that the development of precise tables comprising detailed descriptions of the rating indicators is still a subject of research for the DMEA. The DCI does not replace detailed quantitative risk models, but it provides an actionable first-level indicator that enables companies to prioritize mitigation efforts efficiently. However, the presented method is a straightforward and easily applicable methodology for the identification of potential decoupling-related risks and a corresponding quantification of the risk. The subsequent chapter presents empirical insights into the application of the DMEA across several companies.

5. Empirical Application

The DMEA was developed and tested in collaboration with three German SMEs and one large manufacturing corporation in the metal industry. These organizations have different levels of vertical integration, ranging from individual part production to full machine assembly. Two of these organizations jointly form a small value chain. To ensure that the DMEA method is applicable across diverse industrial contexts, we developed three use cases based on the operational characteristics and requirements of these companies.

The first use case focuses on regional value chains, which are characterized by a wide range of interdependencies between organizations in the upstream and downstream value chains. For example, customers may determine sourcing requirements, while at the same time being strongly dependent on the products delivered by the organization. Although these value chains are regionally structured, some critical components may still originate from non-EU suppliers or overseas markets. The second use case addresses the complex material situation in metal processing, which is shaped by multiple dependencies, including supplier reliability, customer pressure - such as those imposed by original equipment manufacturers (OEMs) - and volatile input markets. Material-intensive production significantly increases vulnerability to price fluctuations, currency risks, and sudden changes in raw material availability. In addition, companies must comply with strict and frequently changing regulatory requirements. Finally, the last two organizations are part of a complex value chain with mutual dependencies arising from their respective roles as customer and supplier. This perspective enables a broader extrapolation of the results to upstream and downstream organizations and thus validates the DMEA approach in a broader context.

In structured interviews with process managers and executives within the organization, we explored and discussed the potential decoupling effects that could lead to the previously identified process deviations and decoupling damage. The causal chain was then narrowed down to external origins in up to two steps.

This procedure was applied consistently across all four organizations. Afterwards, we observed that many external failure causes could be narrowed down to a few factors, including supplier delivery capability, shortages of qualified personnel due to labor market constraints, regulatory requirements related to technology and data availability. For each of the four organizations, executives and process managers rated the potential damage that could be caused by each risk, the probability of detection, the occurrence velocity at which the risk would affect the processes, and the occurrence probability, all on a scale from 1 (minimum) to 10 (maximum). These assessment values were then used to calculate the DCI as described in the previous chapter.

6. Conclusion and Discussion

The DMEA method presented in this paper is suitable for systematically identifying, quantifying, and prioritizing decoupling-related risks in supply chain networks. This method enables organizations to identify weak points in their value-creation processes and to document that crucial supply-chain-related risks have been considered. Similar to the well-established FMEA, the DMEA uses multiple sub-factors to calculate a risk indicator. However, the two methods differ fundamentally in their scope and the structure of the indicator. While the FMEA focuses on risks within a single organization, the DMEA explicitly targets risks that emerge from inter-organizational dependencies and the dynamics of supply chain interactions.

As outlined in Section 4.3, the calculation of the DCI differs from the calculation of the RPN. The proposed formula for the DCI emphasizes the contribution of severity and detection probability and while still accounting for the dynamic nature of external disturbances through occurrence probability and occurrence velocity. For decoupling-related risks with high DCI values, appropriate decoupling strategies can be derived to reduce either the occurrence probability, the detectability probability, or both.

Future adaptations of the DMEA spreadsheet should consider whether the factors of severity and velocity of the impact of a risk can be decreased by actions. The current state implies that only the factors of detection probability and the occurrence probability of risks can be influenced. Depending on the formulation of the decoupling-related risk, the actions might also be able to change the factors of severity and velocity. For example, if the risk describes the shortage of supplied parts in general, accrediting backup suppliers would have a beneficial effect on the factor of occurrence probability, but not on the factor of severity. If the risk in question pertains to the shortage of supplied parts from a particular supplier, the accreditation of backup suppliers capable of delivering equivalent parts to those of the original supplier, will serve to reduce the severity of the risk. However, this measure will not impact the occurrence probability of the risk.

Furthermore, the risks of one supply chain member may propagate to others. Consequently, future research should explore the potential of combining DMEAs of multiple organizations in the same supply chain. Such an integrated perspective could help identify systemic weak points by comparing DCI values for the same risk across different partners. Supporting less robust members of a network by reducing their DCI values would lower the propagation effects and enhance the overall resilience of the network. Collaboration with other supply chain members may also increase the chain transparency and thereby reduce detection uncertainties, for example by revealing additional tiers of suppliers or customers that were previously not visible to the focal company.

Finally, a meaningful comparison of DCI values across organizations requires harmonized definitions and a common understanding of the sub-factors and their rating scales. Establishing such a standardization is a promising direction for the future work to enable cross-company risk assessments and to further validate the DMEA method.

In summary, it can be stated that this paper has established the theoretical basis for using DMEA as a systematic method for identifying and evaluating decoupling-related risks and has applied it to practical use cases for the first time.

Acknowledgements

This work was funded by the German Federal Ministry of Research, Technology, and Space.

References

[1] Ivanov, D., Dolgui, A., 2020. Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. International journal of production research 58 (10), 2904–2915.

[2] Lu, C., Zhang, L., 2021. Managing U.S.-China Technological Decoupling. China Quarterly of International Strategic Studies 07 (04), 409–431.

[3] Pesce, A., 2017. The Decoupling of Emerging Economies: Theoretical and Empirical Puzzle. Journal of Economic Surveys 31 (2), 602–631.

[4] Cha, H., Wu, J., Kotabe, M., 2021. The Vulnerability Problem of Business Ecosystems under Global Decoupling. Management and Organization Review 17 (3), 617–623.

[5] Nguyen, Z., 2024. The Effects of Decoupling and its Implication for Business Executives and Enterprises, in: Óbuda University (Ed.), MEB — 22nd International Conference on Management, Enterprise, Benchmarking. Proceedings II. (MEB 2024), pp. 123–130.

[6] Sheffi, Y., 2007. The resilient enterprise: Overcoming vulnerability for competitive advantage : [with a new preface for the paperback edition, 1. MIT paperback ed. ed. MIT Press, Cambridge, Mass., London, 338 pp.

[7] Felbermayr, G., Krebs, O., 2023. Der volkswirtschaftliche Schaden von Decoupling in Deutschland: Szenarien auf Bundes-, Kreis- und Sektorebene. Stiftung Familienunternehmen, München, 180 pp.

[8] Huang, J., You, J.-X., Liu, H.-C., Song, M.-S., 2020. Failure mode and effect analysis improvement: A systematic literature review and future research agenda. Reliability Engineering & System Safety 199, 106885.

[9] Sharma, K.: Srivastava, S., 2018. Failure mode and effect analysis (FMEA) implementation: a literature review.

[10] Zollondz, H.-D., Ketting, M., Pfundtner, R. (Eds.), 2016. Lexikon Qualitätsmanagement: Handbuch des Modernen Managements auf Basis des Qualitätsmanagements, 2., komplett überarbeitete und erweiterte Auflage ed. De Gruyter Oldenbourg, Berlin, Boston, 1312 pp.

[11] Waser, B.R., Peter, D., Kraner, J., 2025. Prozess- und Operations-Management in Wertschöpfungsnetzwerken: Prozessorientierung, Digitalisierung, Internationalisierung, Nachhaltigkeit der Leistungserstellung, 7., vollständig überarbeitete und wesentlich erweiterte Auflag ed. Versus, Zürich, 1396 pp.

[12] Stamatis, D.H., 2008. Failure mode and effect analysis: FMEA from theory to execution, 2. ed., rev. and expanded, [Nachdr.] ed. ASQ Quality Press, Milwaukee, Wis., 455 pp.

[13] Töpfer, A., 2009. Lean Six Sigma: Erfolgreiche Kombination von Lean Management, Six Sigma und Design for Six Sigma. Springer Berlin Heidelberg, Berlin, Heidelberg.

[14] Wieland, A., Durach, C.F., 2021. Two perspectives on supply chain resilience. J of Business Logistics 42 (3), 315–322.

[15] American Institute of Chemical Engineers, 2008. Guidelines for hazard evaluation procedures, 3rd ed. ed. Wiley-Interscience; CCPS Center for Chemical Process Safety, Hoboken, N.J, New York, N.Y, 542 pp.

[16] International Electrotechnical Commission (IEC), 2006. IEC 61025:2006—Fault Tree Analysis (FTA), 2nd ed. International Electrotechnical Commission, Geneva, Switzerland. https://webstore.iec.ch/en/publication/4311. Accessed 12 November 2025.

[17] International Electrotechnical Commission (IEC), 2016. IEC 61882:2016—Hazard and operability studies (HAZOP studies) — Application guide, 2nd ed. International Electrotechnical Commission, Geneva, Switzerland. https://webstore.iec.ch/en/publication/24321. Accessed 12 November 2025.

[18] Leveson, N.G., Leveson, N., 2011. Engineering a safer world: Systems thinking applied to safety. MIT Press, Cambridge, Mass., 534 pp.

[19] Simchi-Levi, D., Schmidt, W., Wei, Y., 2014. From superstorms to factory fires: Managing unpredictable supply chain disruptions. Harvard Business Review 92 (1-2), 96–101.

[20] Ivanov, D., Dolgui, A., Sokolov, B., 2025. Ripple effect in the supply chain: Definitions, frameworks and future research perspectives, in: , Handbook of ripple effects in the supply chain. Springer, pp. 1–34.

[21] Hosseini, S., Ivanov, D., Dolgui, A., 2019. Review of quantitative methods for supply chain resilience analysis. Transportation research part E: logistics and transportation review 125, 285–307.

[22] Snyder, L.V., Atan, Z., Peng, P., Rong, Y., Schmitt, A.J., Sinsoysal, B., 2016. OR/MS models for supply chain disruptions: A review. Iie transactions 48 (2), 89–109.

[23] Ho, W., Zheng, T., Yildiz, H., Talluri, S., 2015. Supply chain risk management: a literature review. International journal of production research 53 (16), 5031–5069.

[24] Liu, H.-C., Liu, L., Liu, N., 2013. Risk evaluation approaches in failure mode and effects analysis: A literature review. Expert Systems with Applications 40 (2), 828–838.

[25] International Organization for Standardization (ISO), 2019. ISO 22301:2019—Security and resilience—Business continuity management systems—Requirements, 2nd ed. International Organization for Standardization, Geneva, Switzerland 03.100.01; 03.100.70. https://www.iso.org/standard/75106.html. Accessed 12 November 2025.

[26] International Organization for Standardization (ISO), 2019. ISO 31010:2019—Risk management—Risk assessment techniques, 2nd ed. International Organization for Standardization, Geneva, Switzerland 03.100.01; 03.100.70. https://www.iso.org/standard/72140.html. Accessed 12 November 2025.

[27] Caldara, D., Iacoviello, M., 2022. Measuring geopolitical risk. American economic review 112 (4), 1194–1225.

[28] Rosing, M. von, White, S., Cummins, F., Man, H. de, 2015. Business Process Model and Notation—BPMN, in: Rosing, M. von, Scheer, A.-W., Scheel, H. von (Eds.), The complete Business process handbook ; Volume 1. Morgan Kaufmann/Elsevier, Waltham, Mass., pp. 433–457.

Biography

Florian Pleier holds a position as a research associate at the Department of Digital Manufacturing at the University of Augsburg since 2024. His research focuses on the resilience and robustness of production systems.

Simon Rekers is Managing Director of DX Facture GmbH. His research and work focus on resilience, sustainability and digitalization in production industry.

Andreas Röder has been a member of IPRI - International Performance Research Institute since 2023. His research examines how technological progress and algorithmic systems affect sustainability and resilience in value creation networks and their consequences for managerial decision-making.

Daniel Beck is a research associate at the Department of Digitalization and AI in Production at the Fraunhofer Research Institution for Casting, Composite and Processing Technology.

Ben Lüskow is a research associate at the Institute of Mechanical Engineering at the University of Bremen. His work involves simulating material flows and logistics with the aim of building resilient supply chains.

Kirsten Tracht is head of the Bremen Institute of Mechanical Engineering (bime) and professor at the faculty of production engineering at the University of Bremen. She is a founding member of the Scientific Society of Assembly Handling and Industrial Robotics (WG MHI).

Johannes Schilp is head of the Chair of Digital Manufacturing at the University of Augsburg. He is also head of the Processing Department at the Fraunhofer Research Institution for Casting, Composite and Processing Technology and president of the Scientific Society of Assembly Handling and Industrial Robotics (WG MHI).

8th Conference on Production Systems and Logistics

A System Readiness Assessment Of LCA Applications Within The Novel Framework Of Green Digital Twins

Gerald Feichtinger, Wolfgang Posch
Chair of Economics and Business Management, Technical University of Leoben, Austria

Abstract

This contribution focuses on a system readiness assessment (SRA) of life cycle assessment (LCA) based applications within the novel framework of Green Digital Twins (GDT). System-based approaches, such as SRA, help decision-makers improve performance management, particularly for emerging and innovative systems, and identify program/system-related and technical risks in a targeted manner. The proposed conceptual framework of GDTs integrates three key components to enable automated and integrated assessments of life-cycle-based impacts in the future: the idea of digital twins (DT) at the management level, the proven tool of LCAs, as well as the consideration of relational structures in the form of an industrial metaverse (IM). However, the article provides a brief overview of the existing literature, particularly in the field of SRA, at the beginning. Subsequently, a multi-stage SRA for the novel GDT systems is conducted, which requires the derivation of system readiness levels (SRL) by incorporating associated technology readiness levels (TRL) and integration readiness levels (IRL) of the technologies and subsystems under consideration. While TRLs indicate the maturity of individual technologies, IRLs reflect the degree of maturity of the integration of different technologies with each other, and SRLs describe the maturity level of integrated system readiness. However, the presented SRA results indicate that the novel GDT system under investigation attains an SRL value of '3', which corresponds to the SRL state 'Prototyped'. An assessment of selected LCA applications within the overall GDT system yields SRL values ranging from 4 ('Defined') to 1 ('Conceptual'). This initial SRA-based evaluation of the novel GDT system provides a valuable foundation for guiding its further development. The outcomes highlight that system readiness depends not only on the maturity of individual technologies, but also on the effective integration of the different technologies within the overall system.

Keywords

System Readiness Assessment; System Readiness Level; Technology Readiness Level; Integration Readiness Level; Digital Twin; Life Cycle Assessment; Industrial Metaverse; Green Digital Twin.

1. Introduction

Modern engineering and technical systems are becoming increasingly complex, characterized by numerous interacting components and technologies and strong interdependencies between technological and organizational levels. For such complex environments, traditional assessment methods that focus on isolated assessment of the readiness of technologies are insufficient for determining whether the system under consideration is ready for real-world deployment. A helpful and supporting approach for analyzing complex systems are System Readiness Assessments (SRA) to assess both, the maturity of individual technologies used and the robustness of the integration of the technologies in the system under consideration.

DOI: https://doi.org/10.15488/20978
ISSN: 2701-6277

The novel approach of a Green Digital Twin (GDT) exemplifies such a complex system by combining three different but complex dimensions into a unified framework [1,2]: (1) the maturity of a digital twin (DT) based system [3], (2) the scopes of a life cycle assessment (LCA) [5,4,6], and (3) the scope of relational structures based on the fundamental idea of the industrial metaverse (IM) [11,8,7,10,9]. Within this GDT framework, the three dimensions differ in their number of characteristics: The DT maturity dimension comprises five different characteristics (digital model, digital shadow, DT, cognitive DT and federated DT). The LCA dimension includes four scopes (gate-to-gate, cradle-to-gate, cradle-to-grave and cradle-to-cradle). Finally, the relational dimension consists of four different characteristics (product, process, organization and business ecosystem). Considering all characteristics of the three GDT dimensions, a total of 80 distinct system combinations[1] can be identified within the framework. Figure 1 provides an indicative illustration for the GDT system, highlighting one exemplary combination in the form of a 3D cube.

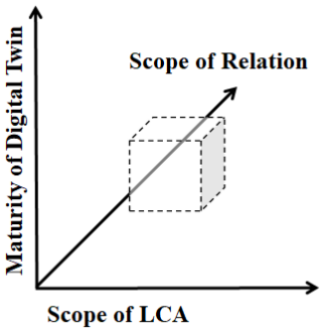


Figure 1: Schematic representation of the Green Digital Twin (GDT) cube (based on [1,2]; own illustration).

Given the novel GDT system and its three different dimensions, assessing the overall system readiness requires understanding not only the maturity of the underlying technologies, but also about the effectiveness with which these technologies interact. Accordingly, this article presents an initial SRA for the proposed GDT concept, analysing both the technological and integrative readiness of the considered subsystems. The complexity from integrating the three previously introduced dimensions constitutes the central motivation for conducting an SRA.

First, a brief overview of the most recent literature in the field of SRA is presented, followed by an introduction to the fundamental methodological approach for the subsequent analyses. Following an overview of the most important SRA results, the weaknesses of the approach are discussed. The paper concludes with a summary and an outlook on future research.

2. Literature review

A concise overview of readiness level methodologies is provided, followed by a review of practical applications of System Readiness Levels (SRL) through SRA.

2.1 Readiness Levels in a Nutshell

Continuous technological innovation requires well-thought-out coordination of various independent components and subsystems. Due to the increasing complexity of the emerging systems, transparent and well-structured approaches are becoming indispensable for assessing technological readiness and existing risks. Readiness levels[2] have emerged as one of the most widely adopted mechanisms, which have been first

[1] Five options for DT maturity, four options for LCA scope, and four options for relation scope: $5 \times 4 \times 4 = 80$.
[2] A detailed list of the readiness levels used in this article is provided in the appendix.

introduced by NASA in the 1990s as *Technology Readiness Levels* (TRL). The TRL approach provides a nine-level scale to evaluate purely technological readiness beginning with the observation of basic principles (level 1) to fully functional deployment in operational environments (level 9). Since then, TRL has become a de facto standard for evaluating technological maturity also having been standardized in ISO 16290. [12]

However, modern systems driven by the digital transformation such as circular-economy solutions or data-driven operations cannot solely be evaluated based on individual technologies, but also require simultaneous maturation of other aspects such as processes, integration capabilities or even data quality. As a result, a variety of complementary frameworks[3] have emerged including *Integration Readiness Levels* (IRL) and *System Readiness Levels* (SRL). These approaches extend the fundamental concept of TRL toward a more holistic approach of readiness, since TRL does not take into account the component-level concerns involving integration and interoperability, which are critical to understand the maturity of an entire system. [13,14]

To address the challenge associated with system integration, IRL were introduced as a structured mechanism to assess the maturity of integration points between components within a system. IRL enable a consistent and comparable evaluation of the interaction of subsystems, independent of the TRL associated with the technologies. Typically defined on a 10-level scale, IRL range from no integration activities between components (level 0) to fully proven integration of components in an operational environment (level 9). IRL provide a systematic way to determine whether subsystems can reliably operate together in relevant or operational contexts by capturing the maturity and interoperability of interfaces. [19,16,20,17,18,15]

In a further step, the maturity of entire systems can be assessed using a holistic measure known as SRL by combing two dimensions of system risk: the technological feasibility of critical components, captured through TRL, and interoperability of the system as a whole, represented by IRL. The integration of TRL and IRL enables SRL to characterize system readiness comprehensively, overcoming the limitations of assessing components or interfaces in isolation. SRL, which have been introduced in the mid-2000s, are typically assessed on a nine-level scale, starting with early system concepts (level 1) to demonstrations of full systems in operational environments (level 9). In this way, the SRL concept addresses the inherent interdependencies within complex systems and acknowledges that a system is only as ready as its least mature component or its most critical integration point. [23–25,21,22,15]

2.2 System Readiness Assessments: A Short Review

Since the publication of SRL, System Readiness Assessments (SRA) have become increasingly important in industrial domains with complex systems which require coordinated development of critical technologies and integration architectures. SRA is a structured evaluation process to determine the maturity and risk with a deploying complex system. First and most important applications emerged in the aerospace and defence sectors to assess the maturity and to identify the system risks in mission systems [23,18,26,15,27]. Beyond that, in more recent case studies, SRA have also been applied in industrial automation and manufacturing [28], maritime and subsea systems [29,30], complex systems of emerging technologies [16], advanced fossil energy applications [31,32], decision support systems [33], farming system transformation [34], but also for cost-minimizing models [35] or carbon optimization in buildings [36]. In many of the applications mentioned, different system-related requirements are examined using the SRA approach, in particular the identification of critical path components during system integration in complex system architectures. In addition, the fundamental mathematical properties and its limitations of the general SRA approach have also already been assessed in detail [37] as well as enhanced assessment approaches [39,38,14].

To date, the authors are not aware of any SRA-based studies on the topic addressed in this article, namely LCA that are embedded in a comprehensive digital system, such as a DT, and go beyond the complexity of

[3] Other frameworks such as Manufacturing Readiness Levels (MRL), Data Readiness Levels (DRL) or Organisational Readiness Levels (ORL) go beyond the scope of work.

a single company, addressed by the IM in terms of the business ecosystem. This article addresses this gap by conducting an SRA for a novel GDT framework that is still at an early conceptual development stage.

3. Methodology

An SRA process comprises three main steps: defining system boundaries, mapping the system, and performing iterative TRL and IRL assessments to derive SRL. These steps are explained in detail below.

3.1 SRL Calculation Methodology

The applied calculation methodology follows a multi-step procedure and relies on a pairwise matrix-based approach for deriving SRL values [40,41,39,15,27]. In this framework, an SRL is first computed for each element (or subsystem) within the system. After that, an SRL for the overall system is obtained by taking the arithmetic mean of the individual SRL values from the element results. In mathematical terms, the SRL can be expressed as a function of TRL and IRL as follows:

$$SRL = f(TRL, IRL) \tag{1}$$

using a TRL vector with n elements

$$TRL_{nx1} = \begin{Bmatrix} TRL_1 \\ \dots \\ TRL_n \end{Bmatrix} \tag{2}$$

and a symmetric IRL matrix for a system of n elements

$$IRL_{nxn} = \begin{Bmatrix} IRL_{11} & \dots & IRL_{1n} \\ \dots & \dots & \dots \\ IRL_{n1} & \dots & IRL_{nn} \end{Bmatrix}. \tag{3}$$

An integration between two elements i and j within the system is represented by the element IRL_{ij} in the form of a representative IRL value. The fictious self-integration of an element i (IRL_{ii}) is always assigned a value of 9. Zeros in the IRL matrix indicate that no integration between the respective elements is planned.

In the next step, the TRL vector and the IRL matrix are normalized[4]. The corresponding SRL values are then derived from these normalized values ($\widehat{TRL}$ and $\widehat{IRL}$) via a matrix-vector product as follows:

$$SRL_{nx1} = \widehat{IRL}_{nxn} \times \widehat{TRL}_{nx1} \,. \tag{4}$$

Each individual element SRL_i thereby represents the system readiness of a specific element i with all its connections to neighboring elements in the system under consideration. Formally applies:

$$SRL = \begin{bmatrix} SRL_1 \\ \dots \\ SRL_n \end{bmatrix} = \begin{bmatrix} \widehat{IRL}_{11}\widehat{TRL}_1 & + & \dots & + & \widehat{IRL}_{1n}\widehat{TRL}_n \\ \dots & + & \dots & + & \dots \\ \widehat{IRL}_{n1}\widehat{TRL}_1 & + & \dots & + & \widehat{IRL}_{nn}\widehat{TRL}_n \end{bmatrix}. \tag{5}$$

To obtain the final SRL for each element, the SRL values are also normalized. For that, each individual element SRL_i is divided by its number of its nonzero integrations[5] m_i, which is determined from the IRL

[4] Normalization is performed by dividing by 9, which is the maximum value of TRL and IRL.
[5] Here, m_i is the number of entries > 0 in the corresponding row (or column) of the IRL matrix.

matrix. Each m_i thus represents the number of integrations of an element i with all other elements[6] in the system under consideration. The following therefore applies:

$$SRL_{element,i} = \frac{SRL_i}{m_i} . \quad (6)$$

Each element $SRL_{element,i}$ is thus the SRL resulting from the SRA for an individual element in the system under consideration. The final SRL for the entire system is calculated as the arithmetic mean of the individual SRL values of all elements considered in the system. It follows:

$$SRL_{system} = \frac{\sum_{i=1}^{n}\left(\frac{SRL_i}{m_i}\right)}{n} = \frac{\sum_{i=1}^{n} SRL_{element,i}}{n} \quad (7)$$

with SRL_{system} as the final SRL for the overall system under consideration and n as the number of elements considered in the system.

3.2 SRL Translation Model

By definition, the final SRL values for individual elements ($SRL_{element,i}$) as well as for the entire system (SRL_{system}) lie between 0 and 1 as a result of normalization. This requires a separate SRL translation model to convert the resulting SRL values between 0 and 1 into SRL ratings between 1 and 9. [40]

The development of the SRL translation model requires a multi-stage and iterative approach. As part of the procedure, nine different SRL must be computed for the entire system. For each of the nine calculation steps, all TRL and the IRL entries with a value greater than zero uniformly assigned the same value, beginning with the value 1 (for the first iteration) and incrementally increasing to 9 (for the ninth and last iteration). Subsequently the mean values of each pair of adjacent SRL results (one for each state) are used to define the thresholds in the final conversion table of the SRL translation model. This conversion table enables the mapping of SRL values derived from the calculation model into corresponding SRL ratings. A more detailed explanation for the SRL translation model follows in the next section, using a practical example.

4. Results

This section presents an SRA evaluation of a novel GDT system, including input data preparation and the resulting SRL and SRL translation model.

4.1 Input Data: TRL and IRL

As outlined above, the SRL is determined through a multi-stage procedure. The first step involves selecting the technology options. In the case of the novel GDT, these are a DT system and its associated maturity levels, the LCA tool with its various assessment scopes, and the relation dimension in various extends.

Meaningful TRL are collected for all three technologies and the associated characteristics, which are supplemented by own assumptions. The technological maturity of DT system is very domain-specific and is rated 5 overall, with its characteristics rated between TRL 2 (federated DT) and 8 (digital model). LCAs are generally already very mature and rated at TRL 8, while the various analysis scopes are rated at TRL 6 (cradle-to-cradle) to TRL 9 (gate-to-gate). The novel scope of the relationship is still in the development stage and is estimated at TRL 4, with the various characteristics estimated at TRL 2 (business ecosystem) to TRL 7 (product level). Table 1 illustrates an overview to all used TRL in the GDT framework.

[6] This also includes the integration of the element under consideration with itself.

Table 1: Technology Readiness Levels (TRL) used for the GDT system (own assumptions).

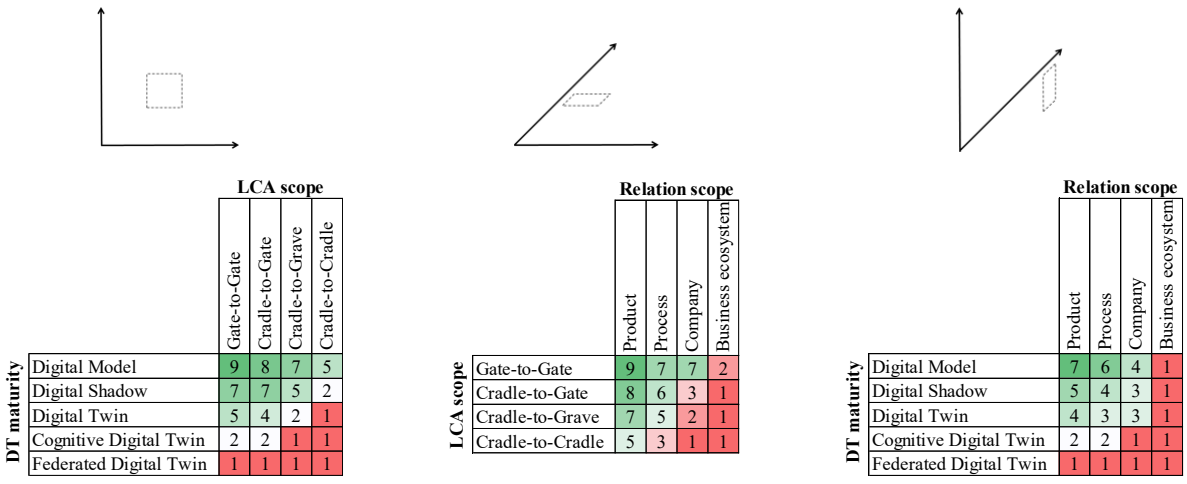
Element	TRL	Characteristic	TRL	Comment
Digital Twin (DT)	5	Digital Model	8	Widely used (7-9)
		Digital Shadow	7	Common (6-8)
	(Domain specific: 2-8)	Digital Twin	5	Mature (4-7)
		Cognitive DT	3	Experimental (2-4)
		Federated DT	2	Early pilots (1-3)
Lifecyle Assessment (LCA)	8	Gate-to-Gate	9	Operational (9)
		Cradle-to-Gate	9	Fully standardized (9)
	(Fully mature: 6-9)	Cradle-to-Grave	8	Very mature (8-9)
		Cradle-to-Cradle	6	Growing (6-7)
Relation Scope	4	Product	7	Widely implemented (6-8)
		Process	6	Multiple pilots (5-7)
	(Early demonstrators: 2-7)	Organization	4	Early demonstrators (3-5)
		Business Ecosystem	2	Largely conceptual (1-3)

For each unique combination – defined by the DT maturity, LCA scope, and relation scope – represented in the GDT cube, a representative TRL is determined using the following formula:

$$TRL = min(TRL_{DT_{level}}, TRL_{LCA_{scope}}, TRL_{REL_{scope}}) \quad (8)$$

As shown, a conservative assessment strategy is adopted, meaning that the lowest TRL associated with the technological maturity of each combination is used.

Following that, individual IRL values for all possible combination options in the GDT cube are determined. To do this, the IRLs for pairwise combinations were first determined: DT maturity with LCA scope, LCA scope with relation scope, and DT maturity with relation scope. Due to the lack of corresponding information in the literature, the selected IRL values are expert estimates made by the authors based on their domain knowledge. Figure 2 provides an overview of the respective combinations, the associated dimensions in the GDT cube, and the selected IRL values.



DT maturity \ LCA scope	Gate-to-Gate	Cradle-to-Gate	Cradle-to-Grave	Cradle-to-Cradle
Digital Model	9	8	7	5
Digital Shadow	7	7	5	2
Digital Twin	5	4	2	1
Cognitive Digital Twin	2	2	1	1
Federated Digital Twin	1	1	1	1

(a)

LCA scope \ Relation scope	Product	Process	Company	Business ecosystem
Gate-to-Gate	9	7	7	2
Cradle-to-Gate	8	6	3	1
Cradle-to-Grave	7	5	2	1
Cradle-to-Cradle	5	3	1	1

(b)

DT maturity \ Relation scope	Product	Process	Company	Business ecosystem
Digital Model	7	6	4	1
Digital Shadow	5	4	3	1
Digital Twin	4	3	3	1
Cognitive Digital Twin	2	2	1	1
Federated Digital Twin	1	1	1	1

(c)

Figure 2: Definition of IRL for the three possible GDT combinations (own assumptions).

As mentioned before, the novel GDT cube consists of 80 distinct system combinations, each of it must be represented in the IRL matrix. Consequently, the final IRL matrix captures all possible system combinations

and therefore comprises 80×80 IRL values[7]. Again, for each unique system combination represented in the GDT cube, a representative IRL is determined using the following formula:

$$IRL = min(IRL_{DT_{level}-LCA_{scope}}, IRL_{LCA_{scope}-REL_{scope}}, IRL_{DT_{level}-REL_{scope}}) \quad (9)$$

Analogous to the TRL procedure, a conservative strategy is adopted for IRL evaluation, in which the minimum IRL of each combination is taken as the representative value.

4.2 System Readiness Levels for GDT

Based on the resulting input data (TRL and IRL), the SRL values and the corresponding SRL translation model are calculated using the methodology described above. To derive the final SRL levels for the novel GDT system, the SRL translation model is constructed by iteratively assigning TRL and IRL values from 1 to 9. This iterative procedure computes nine SRL values, which are then used to subsequently generate upper and lower midpoint estimates that define the boundaries of the SRL value ranges. Table 2 presents the resulting SRL translation model for the GDT system including the SRL values generated during the iterative procedure, the calculated midpoints, and the intervals for each SRL level, together with the abbreviations[8].

Table 2: System readiness level (SRL) translation model for the GDT system (own calculations).

TRL	IRL	SRL system	Midpoint b/w levels[1)]		SRL system range		SRL	SRL
			lower	*upper*	*from*	*to*		*state*
9	*9*	1,0000		0,8975	0,8975	1,0000	*9*	Operational
8	*8*	0,7951	0,7043		0,7043	0,8975	*8*	Interoperable
7	*7*	0,6136		0,5346	0,5346	0,7043	*7*	Demonstrated
6	*6*	0,4556	0,3883		0,3883	0,5346	*6*	Validated
5	*5*	0,3210		0,2654	0,2654	0,3883	*5*	Developed
4	*4*	0,2099	0,1660		0,1660	0,2654	*4*	Defined
3	*3*	0,1222		0,0901	0,0901	0,1660	*3*	Prototyped
2	*2*	0,0580	0,0377		0,0377	0,0901	*2*	Identified
1	*1*	0,0173			0,0000	0,0377	*1*	Conceptual

1) The lower/upper midpoint is calculated as the average of the last two SRL values, respectively.

Subsequently, an SRA-based evaluation of the novel GDT system is conducted to receive the corresponding SRL values, along with the associated SRL levels and SRL states. Based on the SRA procedure, the GDT system obtains an SRL value of '0.0965', which corresponds to SRL level of 3 and an SRL state 'Prototyped'. In addition to assessing the system as a whole, five selected LCA-related applications are analyzed with regard to the integration into this innovative GDT framework.

A product-based LCA using a digital model (averaged over all LCA scopes) reaches an SRL value of '0.2354' and the SRL state 'Defined', representing the most advanced maturity level within the system context. The DT for an organization (Cradle-to-Gate) achieves an SRL value of '0.1086' and the SRL state 'Prototyped', indicating further potential for development. The 'circular economy' approach (cradle-to-cradle) receives an averaged SRL value of '0.0801' and is therefore classified as 'Identified'. More advanced approaches within the GDT framework remain at an early stage: an intelligent DT[9]– represented as the average over all cognitive DT and federated DT combinations – shows an SRL value of '0.0622' with the state 'Identified', while the visionary idea of a *Green Industrial Metaverse* – in the present context defined

[7] The final IRL matrix captures all pairwise combinations of elements considered within the system. For the novel GDT system, transforming the 3D cube into a 2D matrix resulted in a matrix of this size, ensuring that all 80 unique combinations are appropriately represented in the resulting IRL matrix.
[8] A detailed definition of the SRL levels is provided in the appendix.
[9] Represented as in the proposed GDT framework.

as a combination of a federated DT, cradle-to-cradle based LCA, and a business ecosystem – receives an SRL value of '0.0346' corresponding to the SRL state 'Conceptual'. Table 3 summarizes the results.

Table 3. Resulting SRL for the novel GDT system and selected representative system elements (own calculation).

System	SRL (value)	SRL (level)	SRL (state)
GDT cube (entire system)	0,0965	3	Prototyped
Product Life Cycle Assessment (average)	0,2352	4	Defined
Digital Twin for Organization	0,1086	3	Prototyped
Circular Economy (average)	0,0801	2	Identified
Intelligent Digital Twin (average)	0,0622	2	Identified
Green Industrial Metaverse	0,0346	1	Conceptual

5. Discussion

In this article, the established SRA methodology is applied to a novel GDT system, combining DT, LCA, and relational integration (adapted from the IM). The results indicate that the GDT framework is at a prototypical maturity stage, reflecting the authors' expectations. Importantly, the findings should be interpreted in the context of a holistic system evaluation, which shows that even comparatively mature LCA approaches show an unexpectedly low maturity status in the context of the system assessment.

Due to limited empirical data on this emerging GDT concept, several assumptions – particularly in deriving the IRL values, which are central to the SRA– were necessary. These estimates may be subjective and could potentially create a bias on the computed SRL values. A qualitative analysis of how variations in IRL assumptions affect SRL outcomes would provide further insight into the robustness of the results. Consequently, future works could include independent IRL evaluations to develop objective IRL derivation methods and sensitivity analyses to assess the robustness of the resulting SRL values.

With respect to calculating SRL values for the overall system, the literature commonly suggests taking the average across all considered elements. However, this approach may be misleading, and a more conservative 'weakest link' strategy – in which the system's readiness is determined by its least mature component – could provide a more realistic assessment.

Furthermore, the robustness of the SRA results should be investigated in greater detail, potentially through the application of alternative mathematical approaches – for example [39,37,27] – to validate the findings.

6. Conclusion and outlook

The SRA approach proves highly valuable for analyzing the readiness of complex, integrated systems such as this innovative GDT system examined in this article. However, one challenge in this context lies in the accurate and statistically robust definition of the necessary input data—in particular the IRL values. Due to the limited available of empirical data for determining IRL values, practitioners must rely on a certain degree of prior expertise, which introduces uncertainty into the assessment. Nevertheless, the SRA approach provides a systematic approach to guide the further development of the GDT system in structured way.

The presented results provide guidance for further research to advance this prototypical concept. A methodological approach is needed to derive IRL values free from subjective bias, and future work could include a sensitivity analysis to assess result robustness. In addition, a complementary analysis of various application clusters within the GDT cube would be beneficial – an aspect that was only briefly touched in the evaluation of LCA-related sub-elements earlier. Further assessment methods could also be incorporated, such as a Design Structure Matrix approach as discussed in [42,27]. Moreover, applying alternative mathematical methods for calculating the SRL values, for example [27], could help evaluating the robustness of the results.

Appendix

Table 4. Definition of System Readiness Levels (based on [40]; own supplements).

Level	State	Description
9	Operational	System has achieved initial operational capability and can satisfy mission objectives
8	Interoperable	System interoperability should have been demonstrated in an operational environment
7	Demonstrated	System threshold capability should have been demonstrated at operational performance level using operational
6	Validated	System component integrability should have been validated
5	Developed	System high-risk component technology development should have been complete; low-risk system components
4	Defined	System performance specifications and constraints should have been defined and the baseline has been allocated
3	Prototyped	System high-risk immature technologies should have been identified and prototyped
2	Identified	System material solution should have been identified
1	Conceptual	System alternative material solutions should have been considered

Table 5. Definition of Integration Readiness Levels (based on [40]; own supplements).

Level	State	Description
9	Proven	System Integration is proven through successful mission-proven operations capabilities
8	Qualified	System integration completed and mission qualified through test and demonstration in an operational environment
7	Demonstrated	System prototype integration demonstration in an operational high-fidelity environment
6	End-to-end	Validation of interrelated functions between integrating components in a relevant end-to-end environment
5	Relevant	Validation of interrelated functions between integrating components in a relevant environment
4	Lab	Validation of interrelated functions between integrating components in a laboratory environment
3	Design	The detailed integration design has been defined to include all interface details
2	Requirement	There is some level of specificity of requirements to characterize the interaction between components
1	Concept	A high-level concept for integration has been identified
0	Unintegrated	No integration between specified components has been planned or intended

Table 6. Definition of Technology Readiness Levels (based on [12]).

Level	State	Description
9	Proven	Actual system proven in operational environment
8	Qualified	System complete and qualified
7	Operational	System prototype demonstration in operational environment
6	Demonstrator	Technology demonstrated in relevant environment
5	Relevant	Technology validated in relevant environment
4	Lab	Technology validated in lab
3	Proof of concept	Experimental proof of concept
2	Technology	Technology concept formulated
1	Basic	Basic principles observed

References

[1] Feichtinger, G., Mark, D., Schlar, D., Posch, W., 2024. Green Digital Twin for Organizations as Enabler for Carbon Management.

[2] Feichtinger, G., Posch, W., 2026 (forthcoming). Towards Green Digital Twins: A Conceptual Framework Integrating Digitally Enabled Life-Cycle Assessment and Ecosystems.

[3] Liu, Y., Feng, J., Lu, J., Zhou, S., 2024. A review of digital twin capabilities, technologies, and applications based on the maturity model. Advanced Engineering Informatics 62, 102592.

[4] Hauschild, M.Z., Rosenbaum, R.K., Olsen, S.I., 2018. Life cycle assessment: Theory and practice. Springer, Cham, 1216 pp.

[5] Davila R, M.F., Schwark, F., Dawel, L., Pehlken, A., 2023. Sustainability Digital Twin: a tool for the manufacturing industry. Procedia CIRP 116, 143–148.

[6] Mügge, J., Seegrün, A., Hoyer, T.-K., Riedelsheimer, T., Lindow, K., 2024. Digital Twins within the Circular Economy: Literature Review and Concept Presentation. Sustainability 16 (7), 2748.

[7] Kshetri, N., 2023. The Economics of the Industrial Metaverse. IT Prof. 25 (1), 84–88.

[8] HUANG, S., LUO, H., ZHENG, P., MA, N., CHEN, J., MO, G., JING, S., WANG, G., YAN, Y., 2023. Towards Industrial Metaverse: Opportunities and challenges. Journal of advanced manufacturing science and technology 3 (4), 2023011.

[9] Yao, X., Ma, N., Zhang, J., Wang, K., Yang, E., Faccio, M., 2024. Enhancing wisdom manufacturing as industrial metaverse for industry and society 5.0. J Intell Manuf 35 (1), 235–255.

[10] Wirtz, B.W., Langer, P.F., 2025. Industrial Metaverse - Entwicklung und Konzeption der industriellen digitalen Transformation. WIST 54 (3), 11–17.

[11] Hausladen, I., 2025. Das Ökosystem "Industrielles Metaversum" - Fallbeispiel Produktion & Logistik. WIST 54 (5-6), 11–16.

[12] Mankins, J., 1995. Technology Readiness Levels: A White Paper, Washington DC, USA.

[13] Mostafa Ashkavand, Sebastian Jentsch, Natascha Eggers, Torsten Birth, 2023. READINESS LEVELS - Indicators for systems engineering evaluation - An overview.

[14] Tan, W., Sauser, B., Ramirez-Marquez, J., 2011. Analyzing Component Importance in Multifunction Multicapability Systems Developmental Maturity Assessment. IEEE Trans. Eng. Manage. 58 (2), 275–294.

[15] Sauser, B., Verma, D., Ramírez-Márquez, J., Gove, R., 2006. From TRL to SRL : The Concept of Systems Readiness Levels.

[16] Knaggs, M., Harkreader, D., Unione, A., Oelfke, J., Ramsey, J., Keaims, D., Sauser, B., Atwater, B. (Eds.), 2017. Nesting in the evaluation of system readiness for complex systems of emerging technologies. IEEE.

[17] Ramirez-Marquez, J., Sauser, B., 2009. System Development Planning via System Maturity Optimization. IEEE Trans. Eng. Manage. 56 (3), 533–548.

[18] Sauser, B., Gove, R., Forbes, E., Ramirez-Marquez, J., 2010. Integration maturity metrics: Development of an integration readiness level. Information Knowledge Systems Management 9 (1), 17–46.

[19] Jesus, G.T., Chagas Jr., M.F., 2018. Integration Readiness levels Evaluation and Systems Architecture: A Literature Review.

[20] Long, J.M., 2011. Integration readiness levels, in: 2011 IEEE Aerospace Conference. Big Sky, Montana, USA, 5 - 12 March 2011. 2011 IEEE Aerospace Conference, Big Sky, USA. 3/5/2011 - 3/12/2011. IEEE, Piscataway, NJ, pp. 1–9.

[21] Sauser, B., Ramirez-Marquez, J., Nowicki, D., Deshmukh, A., Sarfaraz, M., 2011. Development of Systems Engineering Maturity Models and Management Tools, Fort Belvoir, VA.

[22] Sauser, B., Sarfaraz, M., 2012. System Maturity and Architecture Assessment Methods, Processes, and Tools, Fort Belvoir, VA.

[23] B. Sauser, J. Ramírez-Márquez, Romulo B. Magnaye, Weiping Tan, 2008. A Systems Approach to Expanding the Technology Readiness Level within Defense Acquisition.

[24] Jesus, G.T., Chagas Junior, M.F., 2022. Using Systems Architecture Views to Assess Integration Readiness Levels. IEEE Trans. Eng. Manage. 69 (6), 3902–3912.

[25] Kujawski, E., 2013. Analysis and Critique of the System Readiness Level. IEEE Trans. Syst. Man Cybern, Syst. 43 (4), 979–987.

[26] Sauser, B., Ramirez-Marquez, J., 2009. System Earned Readiness Management for Defense Acquisition.

[27] Tompkins, Z., Grenn, M., Roberts, B., 2020. Improving System Maturity Assessments by Incorporating a Design Structure Matrix. IEEE Trans. Eng. Manage. 67 (1), 122–140.

[28] Sivlen, E., Pirinen, R. (Eds.), 2014. Utilization of the integration readiness level in the context of industrial system projects.

[29] Pirinen, R., 2014. Studies of Integration Readiness Levels: Case Shared Maritime Situational Awareness System. 2014 IEEE Joint Intelligence and Security Informatics Conference.

[30] Yasseri, S.F., 2016. A measure of subsea systems' readiness level. uw tech: int j soc uw tech 33 (4), 215–228.

[31] Knaggs, M., Ramsey, J., Unione, A., Harkreader, D., Oelfke, J., Keairns, D., Bender, W., 2015. Application of Systems Readiness Level Methods in Advanced Fossil Energy Applications. Procedia Computer Science 44, 497–506.

[32] Yasseri, S., Bahai, H., 2018. System Readiness Level Estimation of Oil and Gas Production Systems. ijcoe 2 (2), 31–44.

[33] Lowe, D.C., Justham, L., Everitt, M.J., 2024. Multi-index analysis with readiness levels for decision support in product design. Technological Forecasting and Social Change 206, 123559.

[34] Frija, A., 2024. A system readiness approach to support the packaging and scaling of innovation bundles for farming systems transformation. Agricultural Systems 221, 104148.

[35] Magnaye, R.B., Sauser, B.J., Ramirez-Marquez, J.E., 2010. System development planning using readiness levels in a cost of development minimization model. Systems Engineering 13 (4), 311–323.

[36] Xue, S., Liu, Y., Zhang, B., Guo, X., Zhang, Q., Shi, C., Fan, L., Wang, L., Chang, X., Wang, Y., 2025. Optimizing carbon reduction technologies portfolios of buildings: Lifecycle approach based individual and system technology readiness level assessments. Energy 338, 138832.

[37] McConkie, E., Mazzuchi, T.A., Sarkani, S., Marchette, D., 2013. Mathematical properties of System Readiness Levels. Systems Engineering 16 (4), 391–400.

[38] Tan, W., Ramirez-Marquez, J., Sauser, B., 2011. A probabilistic approach to system maturity assessment. Systems Engineering 14 (3), 279–293.

[39] London, M.A., Holzer, T.H., Eveleigh, T.J., Sarkani, S., 2014. Incidence matrix approach for calculating readiness levels. J. Syst. Sci. Syst. Eng. 23 (4), 377–403.

[40] Austin, M.F., York, D.M., 2015. System Readiness Assessment (SRA) an Illustrative Example. Procedia Computer Science 44, 486–496.

[41] Austin, M.F., York, D.M., 2016. System Readiness Assessment (SRA) a Vade Mecum, in: Auvray, G. (Ed.), Complex Systems Design and Management. Proceedings of the Sixth International Conference on Complex Systems Design and Management, CSD&M 2015. Springer International Publishing AG, Cham, pp. 53–68.

[42] Eppinger, S.D., Browning, T.R., 2012. Design structure matrix methods and applications. MIT Press, Cambridge, Mass., London, 334 pp.

Biography

Gerald Feichtinger (*1980) holds a doctorate in economics and is a senior scientist at the Chair of Economics and Business Administration at the Technical University of Leoben. His interdisciplinary research focuses on solutions for integrated resource and life cycle management in the context of the sustainable and digital transformation of companies and value chains.

Wolfgang Posch (*1970) has been head of the Department of Economics and Business Administration at the Technical University of Leoben since 2021, with a research focus on resource economics, energy management, and business analytics. He spent six years as a management consultant at Arthur D. Little and Accenture, and 15 years as chief strategist and chief economist in the upstream division of the oil and gas company OMV.

8th Conference on Production Systems and Logistics

A Framework For The Knowledge-Based Decision Support System For Deviation Management On The Shop Floor

Yuxi Wang, Joachim Metternich
Institute for Production Management, Technology and Machine Tools, Technical University of Darmstadt, 64287 Darmstadt, Germany

Abstract

Deviation management (DM) is critical for production stability and continuous improvement, yet much shop-floor knowledge remains unstructured, limiting effective decision support. This paper proposes a Knowledge-Based Decision Support System (KB-DSS) for DM. Based on a systematic literature review (SLR), a four-layer framework comprising data, knowledge, application, and user interface layers is designed. The framework integrates heterogeneous deviation-related information into structured, actionable knowledge and links it to real-time deviation events. By leveraging data-science techniques and large language models, the system supports knowledge retrieval, reasoning, and acquisition across the DM workflow. The KB-DSS is implemented and evaluated in multiple industrial environments, demonstrating its scalability and practical applicability for knowledge-driven DM in manufacturing.

Keywords

Deviation Management; Decision Support System; Knowledge Management; Large Language Models

1. Introduction

Deviation management (DM) is a critical factor for maintaining production stability, ensuring product quality, and enabling continuous improvement in manufacturing [1]. A production system is considered to be in a state of deviation when it fails to meet predefined targets derived from standard work or improvement goals [2]. If deviations are not addressed promptly, they can lead to increased downtime, reduced efficiency, and significant value loss [3,4]. Fast and well-informed responses are therefore essential for minimizing these effects [5]. In practice, effective DM relies on both domain expertise and timely access to relevant information. However, much of this knowledge exists in unstructured forms such as operator experience, maintenance logs, or informal documentation. As a result, decision-making often depends on individual expertise rather than systematic support, making DM slow, inconsistent, and difficult to scale [6].

Knowledge-based decision support systems (KB-DSS) provide a promising solution by structuring and contextualizing knowledge, linking it to real-time production data, and guiding users through standardized decision processes [5]. Prior research has already confirmed the potential of knowledge-based systems in key areas such as anomaly detection, root cause analysis, and decision recommendation in manufacturing environments [7,8]. The continuous advancement of technology, specifically the rapid development of Data Science and Large Language Models (LLMs), has injected powerful new possibilities into the evolution of KB-DSS [9]. However, existing solutions often suffer from functional fragmentation, lack of end-to-end integration, and limited intuitive support for expert knowledge interaction [10]. This study addresses these shortcomings by proposing an integrated KB-DSS framework that supports the full lifecycle of DM on the shop floor. It synthesizes current research trends and gaps through a systematic literature review (SLR),

DOI: https://doi.org/10.15488/20979
ISSN: 2701-6277

develops a framework that integrates knowledge representation, real-time data, and intelligent interaction for end-to-end deviation handling, and validates the approach through industrial implementation and evaluation, demonstrating its technical performance, user acceptance, and practical applicability.

The remainder of this paper is structured as follows. Section 2 presents the SLR with key trends and research gaps. Section 3 introduces the framework of the KB-DSS. Section 4 describes the implementation and evaluation approach, and Section 5 concludes the paper and outlines future research directions.

2. Systematic literature review

2.1 Methodology

A SLR was conducted to identify existing approaches, functions, and design patterns for knowledge-driven DM on the shop floor. Following the PRISMA structure, the review included planning, conducting, and reporting phases as illustrated in Figure 1 [11]. The final set of studies was complemented through backward and forward snowballing. The included papers were analyzed with respect to DM objectives, system functions and architectures, and evaluation practices, followed by a qualitative thematic synthesis.

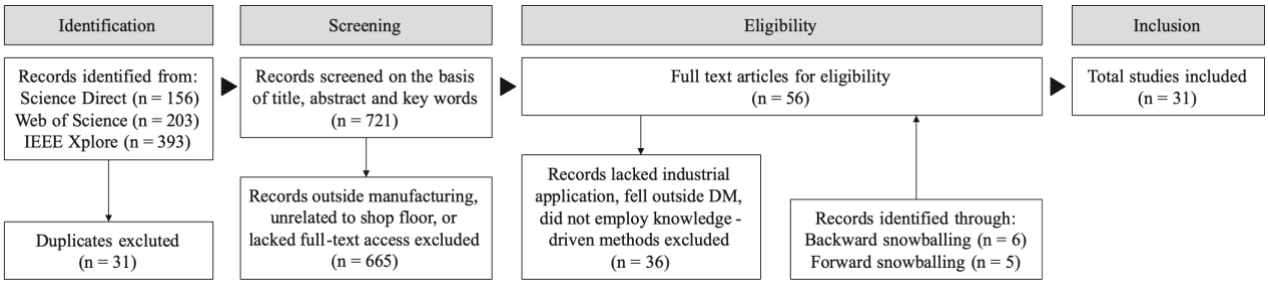


Figure 1 The process of SLR according to PRISMA

2.2 Existing solutions

The requirements of the KB-DSS for DM are derived from challenges identified in the literature and from what previous studies propose as useful decision-support functionalities [12]. The existing solutions in the Table 1 show how current research addresses these requirements.

Table 1: Summary of the core functions in existing literature

DM Phases	Requirements	Existing solutions	References
Deviation Recognition	Early deviation detection	• Rule-based deviation detection with real-time data • Ontology-based reasoning for deviation detection • Deviation prediction with probability	[13–15] [16] [17]
	Structured deviation input	• Standard input formular with defined deviation-related features defined by different methods such as FMEA • Deviation completion with knowledge e.g. ontology • Generation of relevant features or topics for deviation	[18,15,19,20] [21–26] [27–29]
Analysis and Prioritization	Root cause analysis	• Root cause diagnosis with ML and data-based methods • Root cause identification based on semantic reasoning	[13,17,30,23] [18,16,19,27,31,21,22,32–34,29,35–38,20,39,25]
	Decision support for countermeasures	• Countermeasure suggestions with historical records	[18,13,19,40,27,28,41,21,22,32–34,29,35,36,23,38,20,24,39,25,26]
	Prioritization and escalation	• Ranking and categorizing deviations based on weighted scores	[42]
Process Improvement	PDCA-oriented guidance	• Suggestions to the execution and tracking of deviation with historical cases	[19,27,32,29,36,20]
Whole DM process (covering all DM phases)	Automated documentation	• Structured reporting and documentation back to knowledge base.	[18,13–15,19,40,27,41,33,29,23,38,20]
	Real-time dashboards	• Visualization of live deviation status for transparency.	[13,17,14,15,40,27,30,41,21,33,29,42,38,24,39]

The comparison indicates that existing research addresses several core requirements of DM, particularly in early detection, structured deviation input, root cause analysis, and the generation of countermeasure suggestions. The literature also demonstrates support for documentation and transparency through automated

reporting and real-time dashboards. However, relatively few contributions explicitly target prioritization or PDCA-oriented guidance, which leads to uneven coverage across the full DM workflow. Overall, the results suggest that while many individual functions are well supported, integrated solutions that span all phases of DM remain limited. Motivated by this research gap, this study aims to explore a more comprehensive system design that can better align with the overall set of KB-DSS requirements for DM on the shop floor.

2.3 System architecture patterns

The reviewed studies do not follow a single architectural blueprint, but their system elements can be systematically categorized into four layers: data, knowledge, application, and user interface (UI). This categorization allows for a structured comparison of different approaches as shown in Table 2.

Table 2: Summary of the core layers and system elements in existing literature

Core layers	System Elements	References
Data Layer	• Real-time machine & process data • Historical deviation cases and countermeasures • Organisational knowledge	[13,17,14,15,30,31,28,41,37,25] [18,31,28,21,29,35,37,20,39,26] [19,40,35,36]
Knowledge Layer	• Rules, logic & coded knowledge • Ontologies & semantic representations • Knowledge graphs (KGs) • Case-based knowledge	[18,14,15,19,33,23,39] [16,40,30,22,34,37,38] [21,34,29,25,26] [19,27,31,28,20]
Application Layer	• Rule-based and logic-driven reasoning • Case-based reasoning (CBR) • Semantic and graph-based reasoning including LLM-based methods • Machine learning (ML) and deep learning	[18,13–15,30,23] [19,20,24] [17,40,27,28,21,22,32,34,29,35–38,24,25] [13,17,41,23,25,26]
UI	• Real-time monitoring and system awareness interfaces • Diagnostic explanation and decision-support interfaces • Conversational assistants and chat-based troubleshooting • Knowledge capture, reporting, and collaboration tools	[13,17,30,41] [18,14,21,22,32,29,20,25] [19,27,28,34–36] [40,32]

Overall, the findings reveal several recurring trends and insights that suggest a possible architectural direction for designing industrial decision-support systems. The results show a consistent dependency on diverse operational data sources and a growing trend toward using semantically rich knowledge models to structure this data. An increased integration of advanced reasoning methods, including both ML and LLMs techniques, is also observed. In line with these observations, the development pathway of this study is also defined by the transformation of raw operational data into formal knowledge models, which are processed using appropriate inference or learning modules, and delivered via UIs that ensure the clear and context-aware decision supports for shop floor workers and managers.

2.4 Evaluation practices

The aim of this analysis is to identify suitable evaluation methods for the planned study by examining how existing research validates industrial decision-support systems. Based on the review, the evaluation approaches used in the literature can be organised into four categories, reflecting the distinct levels at which these systems are assessed as shown in Table 3. Analytical verification indicates the internal consistency and conceptual soundness of the artifact, which appears in all reviewed works and is therefore omitted from the table, as it offers no basis for differentiation. Existing studies predominantly showcase strong algorithmic performance via quantitative technical testing, using metrics like accuracy, ranking quality, and specific decision-support indicators. Furthermore, practical feasibility is confirmed in several works through deployments in industrial settings or detailed case studies.

Table 3: Summary of the evaluation concepts in existing literature

Evaluation methods	Description	Activities	References
Technical Performance Evaluation	Technical performance and robustness of algorithms and system functions	• Model accuracy and error-based performance measures • Retrieval, ranking, and similarity-quality measures • Decision-support effectiveness and operational quality metrics	[18,17,30,41,32,29,37,24,39] [28,21,32,24,26] [28,33,36,42,38]

Field deployment and case studies	Real-world applicability and workflow integration	•	Real-world deployment in operational production environments	[15,29,23,38,20,39]
		•	Case studies, prototype demonstrations, and mock-up	[14,19,40,31,28,22,35]
Human-centered evaluation	Usability, acceptance, perceived value, and actual task performance	•	Usability and user interaction assessments	[40,27,36]
		•	Expert or focus group interviews	[40,34,36]

However, human-centered assessments are notably scarce. While these evaluations, which address usability, acceptance, and operator task performance, offer critical practical insights, they are not the norm. Since technical metrics can only guarantee system reliability and not real-world applicability, the results suggest that an effective evaluation strategy for this study must integrate rigorous technical testing with user-oriented assessment methods to ensure both functional robustness and practical, operational impact.

3. Framework of the KB-DSS for DM on the shop floor

Based on insights from the SLR, this paper proposes a four-layer KB-DSS framework for DM on the shop floor, as illustrated in Figure 2. The framework follows architectural patterns identified in prior work, emphasizing reliable data handling and semantically organized knowledge as the foundation for knowledge-based decision support. At the application layer, CBR and graph-based reasoning are integrated within the semantic reasoning module to support context-aware deviation analysis, while ML-based approaches are intentionally excluded to maintain a clear focus on knowledge-driven decision support. Given the limited availability of structured operational knowledge in real production settings, a dedicated knowledge acquisition module is introduced to continuously enrich the knowledge base with shop-floor input. The UI layer is designed to combine functionalities observed in existing systems with the specific requirements of DM-oriented KB-DSSs, while the details of each layer are discussed in the following subsections.

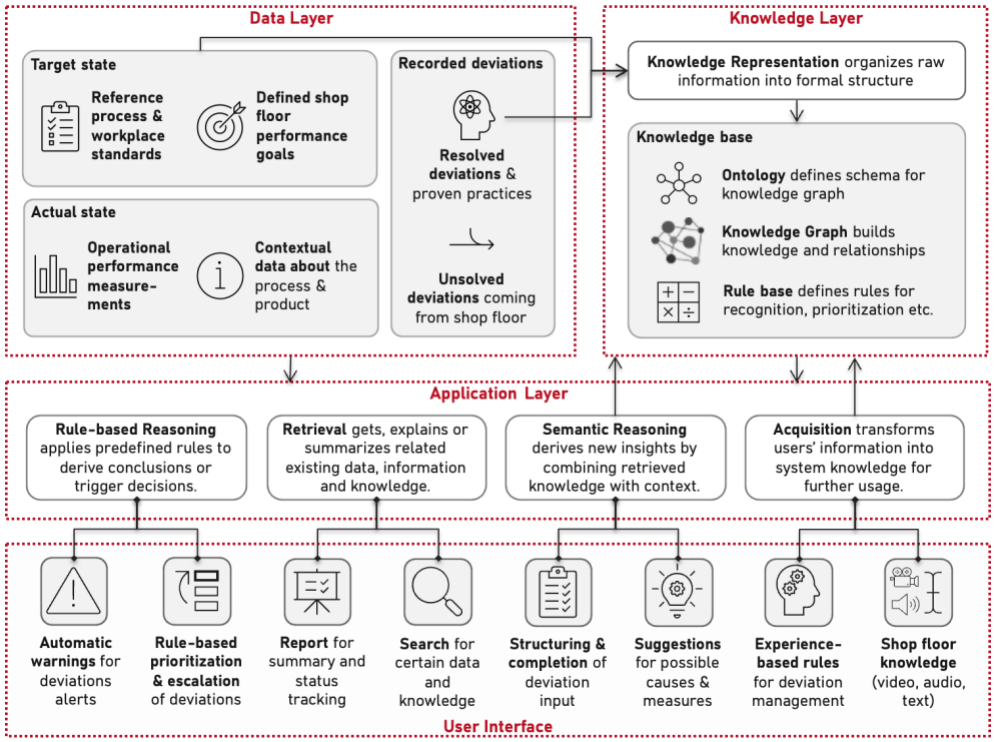


Figure 2 The framework of the KB-DSS for DM on the shop floor

3.1 Data layer

The data collected by the data layer from the production environment to support the DM process can be divided into three categories: target state, actual state, and recorded deviations. Target-state data describe reference process and workplace standards as well as defined shop-floor performance goals. As validated and reusable knowledge, they are extracted into the knowledge representation to support knowledge services

in the application layer. In contrast, actual-state data represent the real-time operational status of the shop floor. They are compared with the target state to detect deviations and to support subsequent decision-making and include operational performance measurements and contextual information related to the current process and product. Recorded deviations comprise both previously resolved deviations with proven practices and unresolved deviations from the shop floor. Resolved deviations are incorporated into the knowledge representation, while unresolved deviations remain linked to the current state. Together, these data categories support understanding deviation context, root-cause identification, and actionable decision support.

3.2 Knowledge layer

The knowledge layer consists of two components: knowledge representation, which transforms target-state data into structured knowledge, and the knowledge base, which supports the knowledge services in the application layer. As shown in Figure 3, knowledge representation serves as a critical bridge between the data layer and the knowledge layer. Its role is to convert heterogeneous data from different sources into structured, semantically rich formats. To connect these data to one knowledge base, a domain ontology for DM is required as the knowledge schema. Within this framework, structured data can be directly incorporated into the KGs by establishing relationships. For unstructured data, natural language processing techniques such as Named Entity Recognition (NER) can be used to extract domain-specific entities (e.g., deviations, causes, countermeasures) and link them within the KGs. LLMs can also assist in constructing and enriching the KGs. Through this process, the knowledge base encodes both explicit knowledge (e.g., rules and standards) and implicit knowledge (e.g., expert experience) into a unified framework, ensuring semantic consistency and interoperability across system components.

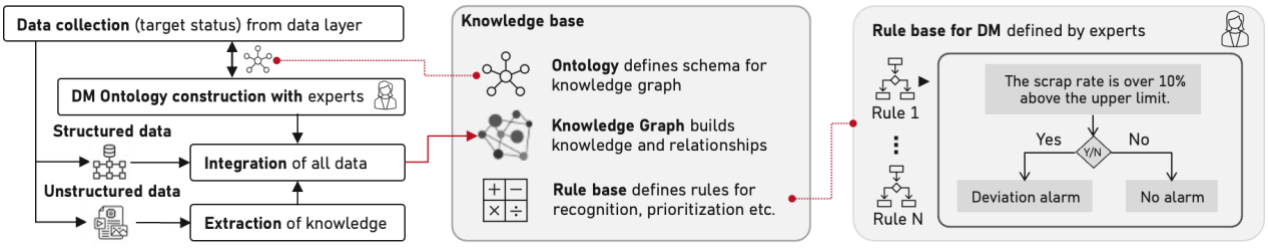


Figure 3 Elements in knowledge layer

Knowledge base integrates the ontology, KG, and rule base. The ontology provides a shared semantic framework that defines concepts, properties, and relationships essential for constructing the KGs. This well-defined structure not only enables the application layer to retrieve relevant knowledge more effectively, but also facilitates the automated generation and maintenance of the KGs [43]. The KG serves as the primary carrier of knowledge. The main graph encapsulates background knowledge related to the shop floor such as production processes, machines, and responsible personnel interconnected through various relations. The subgraph contains more domain-specific knowledge. Historical deviation records can be represented through entities like deviations, root causes, and countermeasures, which are further linked to relevant processes or equipment in the main graph. This enables context-aware root cause analysis and countermeasure suggestion. The rule base encodes experts' operational knowledge into formalized conditional rules and decision logic, allowing these rules to be repeatedly and automatically applied in the application layer.

3.3 Application layer

The objective of the application layer is to connect the data layer, the knowledge layer, and the UI. As illustrated in Figure 2, the application layer is divided into four modules. The first module is retrieval, which focuses on obtaining and presenting content from both the data layer and the knowledge layer. Data access within the data layer can be achieved through direct queries to specific databases. Accessing knowledge layer relies on Cypher queries for fine-grained retrieval from the KGs [44]. LLMs support knowledge retrieval by translating natural language inputs into KG-based queries, generating Cypher queries, and

converting retrieved graph content into user-oriented responses, including summarization and translation [43]. The second module is rule-based reasoning, which infers conclusions by combining rule knowledge with real-time system data. The rule-based reasoning mechanism incorporates both simple if-then rules such as "if the scrap rate exceeds 5% for three consecutive days, a deviation alert is automatically generated", and more advanced analytical methods, such as computing a self-adjusted error score to prioritize deviations [6]. The third module performs semantic reasoning, whose main objective is to integrate semantic knowledge from the knowledge base with data from the current system state, supplemented by additional user inputs, to provide context-specific decision support. In practice, Retrieval-Augmented Generation (RAG) frameworks are widely used for semantic reasoning [45]. The framework adopts a knowledge-graph-based RAG approach for retrieval and response generation, enabling actionable insights such as root-cause identification and countermeasure suggestions [46]. The final module concerns knowledge acquisition, which enriches the knowledge layer and strengthens retrieval and reasoning applications. Knowledge acquisition can be achieved through multiple channels. The first channel is in-process knowledge acquisition, in which information related to newly resolved deviations is extracted and stored in the knowledge base. The data collected through this method are typically well-structured and can be integrated into the knowledge base with minimal preprocessing. The second channel involves conducting dedicated knowledge-capture activities with shop-floor experts, using text, speech, or video to document their experience [5]. LLMs are used to transform multimodal expert inputs into structured knowledge representations [45]. The use of Large Vision-Language Models (LVLMs) further enables the extraction of visual information such as shop-floor layout or details of operational workflows, thereby further expanding the knowledge base of the KB-DSS.

3.4 User Interface

The UI functions as the primary human-computer interaction, mediating between users and the underlying application, data, and knowledge layers. It translates analytical outputs from the four core application modules into comprehensible and actionable forms.

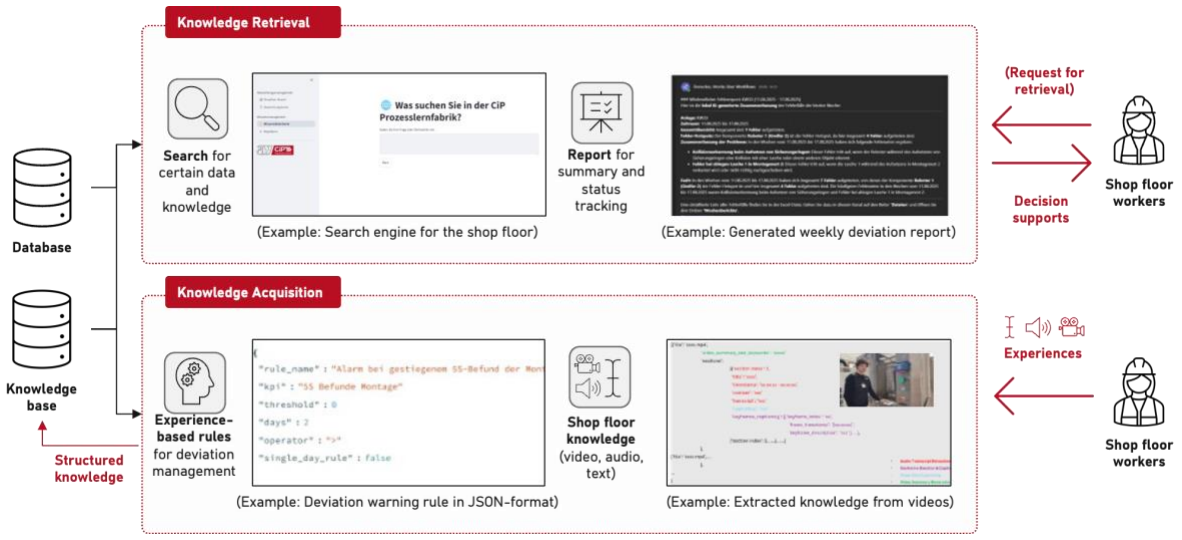


Figure 4 Knowledge retrieval and acquisition with developed demonstrators

Figure 4 illustrates how the knowledge retrieval and acquisition modules interact with both the data layer and the knowledge layer to provide decision support for shop-floor users. The simplest form of knowledge retrieval occurs when users directly search in the knowledge base, such as checking the status of an ongoing deviation or consulting standard operating procedures. These actions can be executed through straightforward database queries. More complex interactions involve generating summaries or reports related to DM, which require additional processing of the retrieved data and knowledge with LLMs. In contrast, the knowledge acquisition module collects shop-floor experience from deviation-handling activities and converts it into structured knowledge. Additional worker experience captured in multimodal formats such

as text, audio, or video is also gathered and incorporated into the system. These functionalities collectively ensure that users can access relevant knowledge on demand, while the system continuously integrates new on-site experience to expand and refine the knowledge layer.

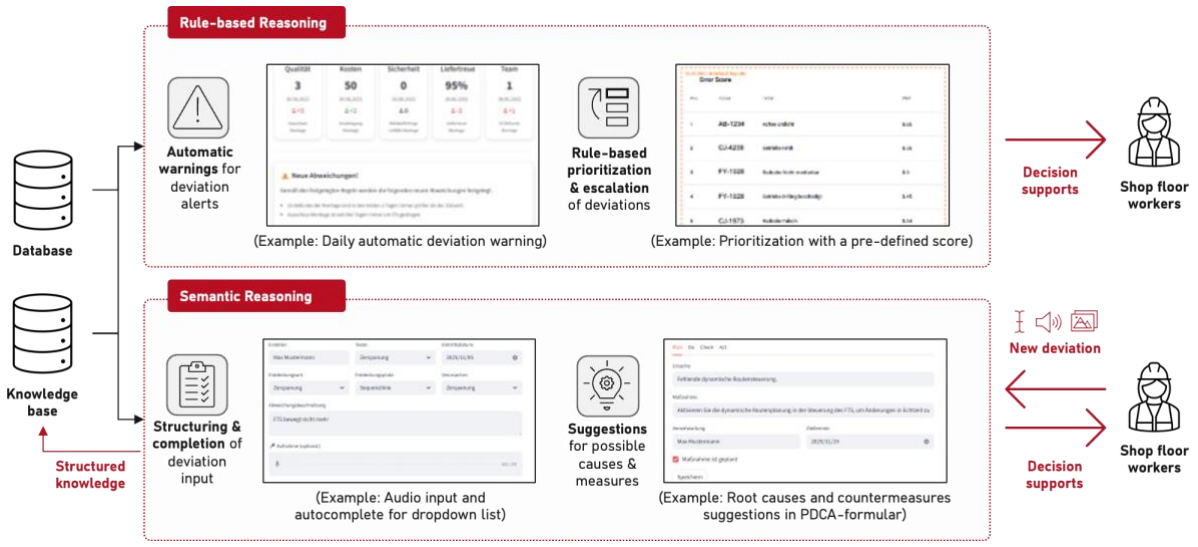


Figure 5 Rule-based reasoning and semantic reasoning with developed demonstrators

Figure 5 illustrates how the rule-based and semantic reasoning modules provide decision support for workers in DM. The rule-based reasoning module automatically processes production data to enable deviation alerting, prioritization, and escalation. By leveraging codified rule knowledge, manual data monitoring is reduced, allowing workers to focus on value-adding activities. Moreover, the rules can be transferred across similar shop-floor environments, supporting system scalability. Semantic reasoning offers two primary forms of decision support. First, during deviation description, it assists with structuring and completing user inputs. Workers may describe deviations using speech, and the system automatically identifies relevant contextual information and attempts to complete the deviation description. This significantly reduces the time required for user input while enhancing the quality of deviation documentation, which in turn affects the quality of subsequent deviation analysis. Second, semantic reasoning supports root-cause analysis and countermeasure generation for new deviations. The generated results are automatically populated into PDCA forms, streamlining worker-system interaction and the DM workflow.

4. Implementation and evaluation

4.1 Implementation in Laboratory and Industrial Environments

The proposed framework has been implemented and evaluated in multiple application settings, including the Learning Factory at TU Darmstadt and several partner companies operating production lines in the automotive and machinery sectors. In industrial deployments, the KB-DSS operates entirely on local infrastructure due to the sensitivity of company-specific data and knowledge. The demonstrator is built using Python, with Streamlit providing the UI, Neo4j supporting the KG, Protégé used for ontology development, and Ollama and LangChain enabling LLMs integration.

For knowledge representation, various natural-language processing techniques were applied to handle domain-specific production terminology. KG construction combined traditional NER-based entity extraction with LLM-generated Cypher code, enabling a more automated and scalable graph-building process [46,6]. Extraction from unstructured data was supported by open-source tools such as PyMuPDF for PDF parsing and LLaVA for image-based content understanding [47]. The rule-based reasoning component integrates predefined rules from the rule base and optimization algorithms that update rule weights based on continuous database feedback, thereby improving decision quality over time. Knowledge retrieval and semantic

reasoning are powered by a KG-enhanced RAG architecture to ensure relevance and accuracy in decision support [46,47,45]. Within knowledge acquisition, rule extraction is implemented through text-to-code transformation, while multimodal shop-floor knowledge is captured through speech-to-text models (e.g., OpenAI Whisper) and LVLMs (e.g., Qwen 2.5 LV).

4.2 Evaluation

As discussed in Section 2, the evaluation of a DSS can be carried out from several perspectives. For the technical performance of the KB-DSS, the focus lies on assessing the quality of the decision support produced by the system. Different performance metrics are applied depending on the application component. For retrieval, common metrics used in LLMs evaluation such as Mean Reciprocal Rank (MRR), Precision@k, and related retrieval measures can be applied to quantify retrieval accuracy [46,47]. The effectiveness of rule-based reasoning can also be evaluated through accuracy by partitioning historical data into training and test sets. Alternatively, process-related KPIs (e.g., cost or processing time) can be defined as targets, and simulation can be used to assess how rule-based reasoning contributes to KPI improvement [6]. Semantic reasoning presents a greater challenge, as it deals with newly occurring deviations and typically lacks ground-truth datasets. Therefore, the retrieval and generation stages can be evaluated separately to assess performance [48]. For acquisition, automated evaluation of extracted knowledge can be performed using labelled ground-truth data, often complemented by expert human assessment for nuanced validation [49].

Field deployment, case studies, and human-centered evaluation are also essential for assessing the applicability of the KB-DSS. These methods ensure that the system aligns with real deviation-management workflows and that both operators and managers find it acceptable and useful in practice. Field deployment focuses on proof-of-concept testing in pilot areas, while human-centered evaluation can involve user satisfaction surveys, such as the System Usability Scale (SUS) [50], and expert interviews to gather qualitative feedback. From a long-term perspective, these forms of evaluation are critical for ensuring sustained adoption and continuous improvement of the system. Preliminary evaluation activities have already begun; however, a comprehensive assessment remains ongoing and will be completed in future work.

5. Conclusion and outlook

This work proposed a KB-DSS for DM, integrating ontologies, KGs, rule bases, and LLM-assisted reasoning into a unified semantic framework. The system operationalizes four core applications including retrieval, rule-based reasoning, semantic reasoning, and acquisition to provide deterministic decision support, contextual interpretation, and continuous knowledge enrichment. Implementations in the TU Darmstadt Learning Factory and several industrial partners demonstrated the feasibility of combining structured knowledge and multimodal data for practical shop-floor applications.

Despite the global definition and development of the proposed KB-DSS framework, a major limitation remains since a comprehensive evaluation has not yet been completed. According to the evaluation framework, only the technical performance assessment can be validated using historical data. In comparison, field deployment and human-centered evaluation require extended observation and iterative refinement before reliable results can be obtained, which indicates that these components should be addressed in future research. Future work should focus on expanding the system's data sources and analytical capabilities. On the data layer, the system can incorporate sensor-generated time-series data alongside image or video data and text data, enabling a more comprehensive representation of equipment conditions, operator behavior, and contextual information for improved deviation detection. Causal analysis methods can also be investigated to identify key influencing factors and improve the reliability of root-cause identification. Additional efforts should be devoted to refining deployment and evaluation processes to ensure the system's effectiveness and user acceptance in real operational environments.

Acknowledgements

The authors gratefully acknowledge the financial support of the research project “PrePAIR” (project number 13IK037I) funded by the European Union and supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a decision by the German Bundestag.

References

[1] Hertle, C., Tisch, M., Metternich, J., Abele, E., 2017. Das Darmstädter Shopfloor Management-Modell. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 112 (3), 118–121.

[2] Suzaki, K., 1993. The new shop floor management: Empowering people for continuous improvement. The Free Press, New York.

[3] Peters, R., 2009. Shopfloor Management: Führen am Ort der Wertschöpfung. LOG_X, Ludwigsburg, 160 pp.

[4] Pötters, P., Schindler, P., Leyendecker, B., 2018. Status quo Shopfloor Management. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 113 (7-8), 522–525.

[5] Turban, E., 1995. Decision support and expert systems Management support systems. Prentice-Hall, Inc.

[6] Wang, Y., Longard, L., Hertle, C., Metternich, J., 2023. Beyond Pareto analysis: a decision support model for the prioritization of deviations with Natural Language Processing.

[7] Nguyen, Q.-T., Tran, T.N., Heuchenne, C., Kim, P.T., 2022. Decision support systems for anomaly detection with the applications in smart manufacturing: a survey and perspective. Machine Learning and Probabilistic Graphical Models for Decision Support Systems, 34–61.

[8] Pietsch, D., Matthes, M., Wieland, U., Ihlenfeldt, S., Munkelt, T., 2024. Root Cause Analysis in Industrial Manufacturing: A Scoping Review of Current Research, Challenges and the Promises of AI-Driven Approaches. JMMP 8 (6), 277.

[9] Eigner, E., Händler, T., 2024. Determinants of LLM-assisted Decision-Making, 44 pp. https://arxiv.org/pdf/2402.17385v1.

[10] leveraging-industrial-iot-and-advanced-technologies-for-digital-transformation.

[11] Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ (Clinical research ed.) 339, b2535.

[12] Metternich, J., Müller, M., Hertle, C., Longard, L., Wang, Y., 2025. Digital Daily Management. CRC Press, Boca Raton.

[13] A. Arama, E. Villeneuve, C. Merlo, L. L. Salvado, 2022. An approach of decision support system for drift diagnosis in cyber-physical production systems, in: . 2022 IEEE International Systems Conference (SysCon), pp. 1–7.

[14] Arevalo, F., M. P., C.A., Tahasanul Ibrahim, M., Schwung, A., 2022. Production Assessment Using a Knowledge Transfer Framework and Evidence Theory. IEEE ACCESS 10, 89134–89152.

[15] Arévalo, F., Tito, C., Diprasetya, Schwung, A., IEEE, 2019. Fault Detection Assessment using an extended FMEA and a Rule-based Expert System, in: . 2019 IEEE 17TH INTERNATIONAL CONFERENCE ON INDUSTRIAL INFORMATICS (INDIN), pp. 740–745.

[16] Bayar, N., Darmoul, S., Hajri-Gabouj, S., Pierreval, H., 2016. Using immune designed ontologies to monitor disruptions in manufacturing systems. Computers in Industry 81, 67–81.

[17] A. Esteban, V. Cobilean, R. Nawaratne, 2024. Predictive Maintenance with Large Language Models and Transformer-based Survival Analysis, in: . IECON 2024 - 50th Annual Conference of the IEEE Industrial Electronics Society, pp. 1–6.

[18] A. Al-Najim, A. Al-Amoudi, K. Ooishi, M. Al-Naser, 2022. Intelligent Maintenance Recommender System, in: 2022 7th International Conference on Data Science and Machine Learning Applications (CDMA). 2022 7th International Conference on Data Science and Machine Learning Applications (CDMA), pp. 212–218.

[19] Camarillo, A., Ríos, J., Althoff, K.D., 2019. PLM Applied to Manufacturing Problem Solving: A Case Study at Exide Technologies, in: , PRODUCT LIFECYCLE MANAGEMENT, VOL 4: THE CASE STUDIES, pp. 233–247.

[20] Tönnes, W., 2018. Applying data of historical defects to increase efficiency of rework in assembly. Procedia CIRP 72, 255–260.

[21] G. Michael, J. Peng, M. B. Marinov, J. Ovtcharova, 2021. Research on Fault Diagnosis Expert System of CNC Machine Tool Based on Expert Knowledge, in: . 2021 XXX International Scientific Conference Electronics (ET), pp. 1–4.

[22] Guo, Z., Zhou, D., Yu, D., Zhou, Q., Wu, H., Hao, A., 2024. An ontology-based method for knowledge reuse in the design for maintenance of complex products. Computers in Industry 161, 104124.

[23] Mewada, S., Saroliya, A., Chandramouli, N., Kumar, T.R., Lakshmi, M., Mary, S.S., Jayakumar, M., 2022. Smart Diagnostic Expert System for Defect in Forging Process by Using Machine Learning Process. JOURNAL OF NANOMATERIALS 2022.

[24] Xu, Z.G., Dang, Y.Z., 2021. Solution knowledge mining and recommendation for quality problem-solving. Computers & Industrial Engineering 159.

[25] Y. Xu, J. Tao, T. Peng, N. Zhang, 2023. Low-Voltage Electrical Product Quality Problem-Solving Based on Improved Deep Structured Semantic Model, in: . 2023 IEEE 19th International Conference on Automation Science and Engineering (CASE), pp. 1–6.

[26] Z. Zhang, J. Zhang, 2024. Expert Recommendation System Based on Knowledge Graph, in: . 2024 4th International Conference on Robotics, Automation and Intelligent Control (ICRAIC), pp. 528–532.

[27] Colabianchi, S., Tedeschi, A., Costantino, F., 2023. Human-technology integration with industrial conversational agents: A conceptual architecture and a taxonomy for manufacturing. Journal of Industrial Information Integration 35, 100510.

[28] Du, K., Yang, B., Xie, K., Dong, N., Zhang, Z., Wang, S., Mo, F., 2025. LLM-MANUF: An integrated framework of Fine-Tuning large language models for intelligent Decision-Making in manufacturing. Advanced Engineering Informatics 65, 103263.

[29] Kohl, L., Ansari, F., 2024. A Knowledge Graph-based Learning Assistance Systems for Industrial Maintenance. Procedia CIRP 126, 87–92.

[30] D. Geng, X. Fu, 2020. Research on Fault Diagnosis Mechanism of Production Line Equipment Based on Semantic, in: . 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), pp. 220–223.

[31] Donauer, M., Peças, P., Azevedo, A., 2015. Identifying nonconformity root causes using applied knowledge discovery. Robotics and Computer-Integrated Manufacturing 36, 84–92.

[32] Heredia Álvaro, J.A., Barreda, J.G., 2025. An advanced retrieval-augmented generation system for manufacturing quality control. Advanced Engineering Informatics 64, 103007.

[33] Hoerner, L., Schamberger, M., Bodendorf, F., 2023. Using Tacit Expert Knowledge to Support Shop-floor Operators Through a Knowledge-based Assistance System. Comput Supported Coop Work 32 (1), 55–91.

[34] Kernan Freire, S., Foosherian, M., Wang, C., Niforatos, E., 2023. Harnessing Large Language Models for Cognitive Assistants in Factories, in: Proceedings of the 5th International Conference on Conversational User Interfaces. CUI '23: ACM conference on Conversational User Interfaces, Eindhoven Netherlands. 19 07 2023 21 07 2023. ACM, New York, NY, USA, pp. 1–6.

[35] Lin, C.Y., Tsai, T.H., Tseng, T.L., 2025. Generative AI for Intelligent Manufacturing Virtual Assistants in the Semiconductor Industry. IEEE ROBOTICS AND AUTOMATION LETTERS 10 (4), 4132–4139.

[36] Löwhagen, N., Schwendener, P., Netland, T., 2025. Can a troubleshooting AI assistant improve task performance in industrial contexts? INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH.

[37] Okazaki, S., Kaminishi, K., Fujiu, T., Wang, Y., Ota, J., 2025. Fault Cause Identification across Manufacturing Lines through Ontology-Guided and Process-Aware FMEA Graph Learning with LLMs, 31 pp. https://arxiv.org/pdf/2510.15428v1.

[38] Ragala, Z., Retbi, A., Bennani, S., 2023. Implementing sharing platform based on ontology using a sequential recommender system. IJECE 13 (6), 6754.

[39] Xu, Z., Dang, Y., Munro, P., 2018. Knowledge-driven intelligent quality problem-solving system in the automotive industry. Advanced Engineering Informatics 38, 441–457.

[40] Campatelli, G., Richter, A., Stocker, A., 2016. Participative Knowledge Management to Empower Manufacturing Workers. INTERNATIONAL JOURNAL OF KNOWLEDGE MANAGEMENT 12 (4), 37–50.

[41] Fattahi, S., Azarhoushang, B., Kitzig-Frank, H., 2025. Knowledge-Based Adaptive Design of Experiments (KADoE) for Grinding Process Optimization Using an Expert System in the Context of Industry 4.0. JOURNAL OF MANUFACTURING AND MATERIALS PROCESSING 9 (2).

[42] Psarommatis, F., Kiritsis, D., 2022. A hybrid Decision Support System for automating decision making in the event of defects in the era of Zero Defect Manufacturing. Journal of Industrial Information Integration 26, 100263.

[43] Wangerow, K., Wang, Y., Kukulies, J., Metternich, J., 2025. Semantische Datenmodellierung für Fehlermanagement auf dem Shopfloor. Zeitschrift für wirtschaftlichen Fabrikbetrieb 120 (11), 796–802.

[44] Wang, Y., Bretones Cassoli, B., Metternich, J., 2023. Building A Knowledge Graph From Deviation Documentation For Problem-Solving On The Shop Floor.

[45] Wang, Y., Chytraeus Jan, Metternich, J. (Eds.), 2024. Large Language Models for Deviation Management. Springer, 543-550.

[46] Cicek, E., Wang, Y., Metternich, J., 2025. Enhancing Deviation Management Using Knowledge Graph based Retrieval-Augmented Generation System. Proceedings of CIRP ICME.

[47] Wang, Y., Cai, Y., Chytraeus, J., Zhao, K., Metternich, J., 2025. Retrieval-Augmented Generation-based Decision Support System for Deviation Analysis on the Shop Floor. Proceedings of CIRP ICME.

[48] Yu, H., Gan, A., Zhang, K., Tong, S., Liu, Q., Liu, Z., 2025. Evaluation of Retrieval-Augmented Generation: A Survey 2301 (3), 102–120.

[49] Davenport, T.H., Prusak, L., 1999. Working knowledge: How organizations manage what they know, [Nachdr.] ed. Harvard Business School, Boston, Mass., 199 pp.

[50] Lewis, J.R., 2018. The System Usability Scale: Past, Present, and Future. International Journal of Human–Computer Interaction 34 (7), 577–590.

Biography

Yuxi Wang, M. Sc. (*1995) is a research associate and PhD candidate at the Institute for Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt since 2021. Her research interests include the further development and introduction of knowledge-based decision support systems to shop floor management and problem-solving systems.

Prof. Dr.-Ing. Joachim Metternich (*1968) is head of the Production Management division at the Institute for Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt. He is the spokesman of the medium-sized businesses digital competence center Darmstadt; he has been president of the International Association of Learning Factories (IALF) from 2015 to 2023 and he is scientific partner of the Conference on Production Systems and Logistics (CPSL).

8th Conference on Production Systems and Logistics

LLM-Based Multimodal Prompting For Adaptive Robot Control In Production Systems

Dominik Koch[1,†], Jakob Wolber[1,†], Zhuo Shi[3], Bo Cheng Ji[3] , Lucas Bretz[2], Alexander Geiser[1], Felix Baer[3], Martin Benfer[1], Florian Stamer[4], Gisela Lanza[1,2]

[1]*wbk Institute of Production Science, Karlsruhe Institute of Technology (KIT), Kaiserstr. 12, 76133 Karlsruhe, Germany*
[2]*Global Advanced Manufacturing Institute (GAMI, KIT China Branch), 10 Yueliangwan Road, Suzhou 21523, China*
[3]*Siemens Limited, China, 18 Dongfang park, Jinfang Road, Suzhou 215129, China*
[4]*Leuphana University, Universitätsallee 1, 21335 Lüneburg, Germany*
† Both authors contributed equally

Abstract

Large Language Models (LLMs) open new opportunities for adaptive automation in production systems by enabling robots to interpret human instructions and generate context-aware actions. In contrast to conventional robot programming, which requires expert knowledge and frequent reconfiguration, LLM-based control promises greater flexibility and easier interaction between humans and machines. However, generic LLMs still face major challenges when applied to manufacturing environments, as they lack grounding in real-world perception and may produce infeasible or unsafe actions. This paper presents a laboratory demonstrator that evaluates how different prompting strategies affect the performance of an LLM-controlled pick-and-place robot. The study systematically compares zero-shot and multimodal few-shot prompting, where visual examples such as annotated video frames and image captions are integrated into the LLM input. A dedicated evaluation model with metrics for plan success, action success, and plan optimality is used to quantify system behavior. The experimental results demonstrate that multimodal few-shot prompting significantly improves planning accuracy, robustness, and adaptability compared to a zero-shot baseline. These findings illustrate the potential of LLM-driven control for future intelligent production systems that combine semantic reasoning, multimodal perception, and human-interpretable automation.

Keywords

Large Language Models (LLMs); Adaptive Production Systems; Multimodal Few-Shot Prompting; Intelligent Robot Control; Human-Interpretable Automation

1. Introduction

Recent advancements in natural language processing (NLP) with LLMs like generative pre-trained transformers (GPT) have gained a lot of momentum in 2024. These models generate human-like texts, understand context, and enable meaningful conversations [1]. Alongside the growth of LLMs, prompt engineering has emerged as a technique to extend the capabilities of LLMs [2]. Integrating LLMs into robotics is reshaping human-robot interactions but is not trivial [3]. In conventional industrial robotics, tasks like pick-and-place are typically pre-programmed, which requires a reconfiguration when adjusting to new environments or tasks. Another disadvantage is that in traditional robotics, specialized experts are required, which increases the complexity and cost of deployment. LLM powered robots, on the contrary, can interpret natural language, adapt to real-time conditions, and generate actions based on contextual understanding, reducing the need for constant reprogramming. This shift enables robots to handle dynamic, ambiguous

DOI: https://doi.org/10.15488/20980
ISSN: 2701-6277

situations, making them more flexible and autonomous [4] Current research on LLMs in robotics focuses on task planning and action derivation, but lacks progress in feedback, multimodal input, and generalization [5]. This work focuses on a proof-of-concept environment simulating pick-and-place tasks with wooden blocks, providing a valuable testbed for evaluating how LLMs can control robotic systems in dynamic environments. In experiments, the LLM receives and interprets natural language commands and determines both the appropriate tools and their execution sequence. These tools include parametrized Python functions for gripper control or robot movement, chosen based on multimodal inputs such as annotated video frames or sensor data. By exploring these structured yet flexible tasks, the study investigates how LLMs can contribute to more adaptive robotic systems, particularly in scenarios requiring frequent task adjustments and natural human–robot interaction. Using prompt engineering and multimodal input techniques, the impact of few-shot and multimodal prompting on LLM performance in robotic control is assessed. This work introduces an evaluation model with generic tasks, manipulation families, and performance metrics to evaluate LLM effectiveness in robotics. Performed experiments quantify zero-shot performance and show improvement with multimodal few-shot prompting, particularly using annotated video frames as solution examples. Additionally, these experiments show that image captioning enhances object detection, thereby improving the LLM's task-solving capabilities.

2. Fundamentals and Related Work

2.1 Fundamentals

An LLM generates responses by calculating the most probable next tokens based on the statistical distribution of tokens in the training data [6] Although LLMs are generally capable, they must be adapted for specific tasks or domains. This is done through prompt engineering and/or finetuning. Fine-tuning updates a pre-trained neural network's weights with task-specific data, adjusting the model's parameters, creating specialized models. However, it requires labeled data, time, computational resources, risks overfitting, and is challenging for tasks with changing data or specifications [7]. Prompt engineering designs and refines prompts to get specific responses from LLMs without adjusting the parameters of the model [8]. Prompts are generally given within a template, which is a predefined structure that organizes the input into clearly defined sections [5]. When only the prompt is provided, without additional examples, it is referred to as a "zero-shot". However, when the prompt includes additional input, such as solution examples, it is called a "few-shot" prompt [9]. Despite their strengths, LLMs struggle with simple tasks like factual lookup. This limitation can be overcome by allowing LLMs to use external tools, which extend their capabilities beyond text generation and allow them to perform real-world actions [10]. An external tool can include an application programming interface, a database, hardware components such as a robotic gripper, or a module that integrates multiple tools to handle complex tasks. In this context, LLMs function as a reasoning hub, using their decision-making ability to interact with and access these tools [11].

The following illustrates the LLM agent concept with the example request: *„pick-up the green block"*. The LLM agent receives a user request and available tools, e.g., a *json*-file describing available Python functions to the agent. The agent then creates a plan, a sequence of tool calls. Since the LLM predicts tokens, tool selection is probabilistic. Tool execution can be deterministic or probabilistic, depending on the tool. The tool call returns an outcome that the agent processes before selecting the next tool. The agent loop continues until the agent’s final observation fulfills the initial user request. Vision-language models (VLMs) are a popular type of multi-modal networks in current research [12]. For example, computer vision, like object detection, is crucial in robotics [13], but traditional closed-set systems can only detect predefined categories and need new labeled data to add others [14]. Detecting arbitrary objects specified by human language is called open-set object detection [15]. VLMs merge computer vision and NLP and enable arbitrary objects detection specified by human language input. Among other things, VLMs can take images and text as input

and generate text, such as describing visual content or annotations. They are used for tasks like image captioning, object detection, and visual question-answering [13].

2.2 Related Work

This section summarizes recent approaches for structured prompting for code generation in task planning, grounding in the environment, granular robot actions, retrieval-augmented prompting, and using the Planning-Domain-Definition-Language (PDDL) in robotics task planning with LLMs. However, some publications can belong to multiple categories.

Structured prompting for code generation in task planning: Publications structure prompts using action primitives, system roles or environment information to improve LLM-generated code accuracy [16-20]. Strategies include XML-tagged outputs and corrective dialogues for better task planning and error handling [19], JSON templates [20] or VLM to interpret human demonstration videos and generate robot task planning [21].

Grounding the environment: Grounding robots in the real world and verifying action feasibility through environmental feedback are essential for reliable task execution [22,23]. LLMs use textual descriptions, perception APIs, or vision models to interpret environments and ensure feasibility [24]. Methods such as affordance functions [22], Q-function verification [23], and iterative feedback loops improve plan quality and execution accuracy [25]. Luan et al. employ Q&A-based few-shot prompting supported by VLM object detection; one LLM decomposes user requests, while another refines them into grounded tasks [26]. Yang et al. introduce a re-planning prompt that adapts task plans using environmental feedback via CoT reasoning, while ToT offers no performance benefit and increases latency [27]. C. Wang et al. translate multimodal cues—human behaviour, position, gaze, dialogue, and scene information—into language for LLM-based action coordination [28]. Other VLM approaches focus on replanning [29], or replace separate vision and reasoning systems with a unified multi-agent VLM framework [30].

Granular robot actions: LLMs often focus on high-level planning and invoke a limited repertoire of action primitive and struggle with low-level motion planning [31,32]. To bridge this gap, methods integrate kinematic-aware prompting to generate motion trajectory waypoints [32] or structured action token for a policy network to generate precise movement commands [33]. Other approaches use 3D~value maps for motion planning [31] and visual marks to images to enhance grounding in multimodal models [34]. Brohan et al. highlights the promise of VLMs and LLMs but note that their high-level focus results in separate low-level controllers missing rich semantics. Vision language action models (VLA) are introduced to decode actions from images and language. VLAs are fine-tuned VLMs and can be integrated with LLMs in robotics and improved using techniques like CoT reasoning [35].

Retrieval-augmented prompting: The quality of few-shot examples is crucial [36,37]. Retrieval-augmented prompting selects relevant examples from databases to improve few-shot learning and task execution [2]. Enhancements like k-nearest-neighbor retrieval and logit biases [36], and human feedback [37] improve example selection.

Plannnig-Domain-Definition-Language: PDDL is a standardized language for defining classical planning problems using logical expressions and structures. Combining PDDL with LLMs can enhance the performance of planners by ensuring actions follow a clear, predefined structure. LLMs translate natural language requests into structured PDDL action plans [38-40] and a classical solver to generate accurate plans [41].

3. Approach

3.1 Problem Statement

Generic LLMs cannot produce executable robot task plans without adaptation, typically via prompt engineering or fine-tuning. This work focuses on prompt engineering. Although integrating an LLM agent with external robotic tools is feasible, achieving reliable performance remains challenging. Improvements can come from stronger models, better tools, or tailored prompts, but model and tool development requires significant resources, and heavily customized prompts risk overfitting. Few-shot multimodal examples, such as video demonstrations, may support more efficient generalization, though their use in robotics is still under investigation [5]. This experiment compares zero-shot performance with (multimodal) few-shot prompting.

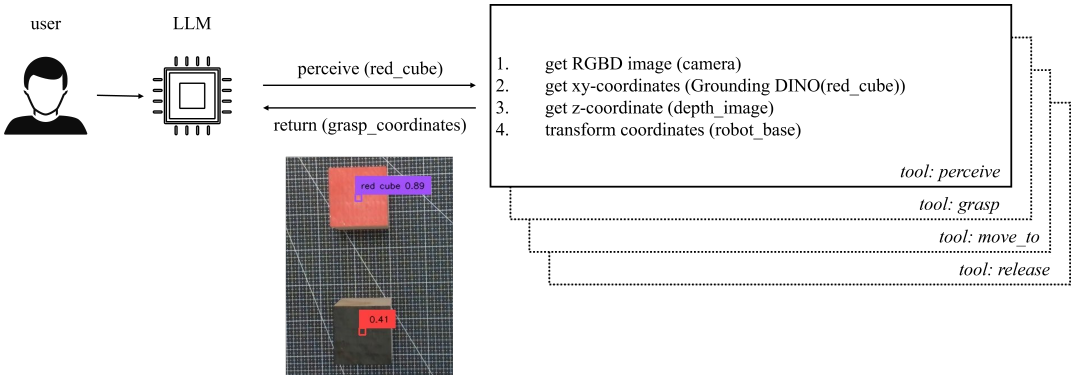


Figure 1: Object detection module: The perceive function uses an RGB-D camera with Grounding-DINO

3.2 Architecture

The proposed approach integrates an LLM agent with external tool execution for robotic control. The system consists of the LLM agent, a UR10e robot, a suction gripper, and an object detection module. Alibaba's Qwen-Max model (temperature 0.95) was used via the Alibaba Cloud API, selected for its strong tool-calling and parameter-selection capabilities [42]. The agent receives a *tools.json* file defining four available tools—*move_to(x,y,z)*, *perceive("object")*, *grasp()*, and *release()*—implemented as Python functions mapped to URScript commands. In this work, "tools" refers to these Python functions, although the term can be broader (Section 2.1). Figure 1 shows the object detection module used by *perceive()*, which outputs the top-center grasp point of the target object. GroundingDINO [15] provides the bounding boxes, and an Intel RealSense D435 RGB-D camera supplies the visual data. For image captioning GPT-4o [43] was used, generating textual descriptions of images [44]. The proposed prompting pipeline consists of six elements: input data (user request), instruction (system role), output indicator (desired format), context (e.g., available tools), solution examples (zero-shot or few-shot), and feedback (re-prompting), following prior work [5].

3.3 Experimental Setup

3.3.1 Metrics

Three metrics are introduced and explained below: Plan success rate (PSR), action success rate (ASR) and optimal plan rate (OPR).

PSR: Defined as the fraction of executions in which the agent produces a correct plan, defined as a tool sequence that reaches the task's goal state. Any additional steps hallucinated after achieving the goal render the plan incorrect. Tool parameters and robot execution are assumed perfect, and plan optimality is not considered. This definition aligns with prior work [25, 45, 46].

ASR: The fraction of executions with correct tool parameters. It evaluates only the agent's selection of Python function parameters under the assumption of flawless tool execution, ignoring external disturbances

such as power loss or camera miscalibration. This separation of plan- and action-level evaluation is common in LLM-based robotics, where execution success is often emphasized [25, 46, 17].

OPR: Is defined as the fraction of executions that generate an optimal plan, defined as using the correct order and the minimum number of tool calls. Non-optimal plans include unnecessary steps or incorrect ordering. Because optimality is stricter than correctness, OPR is always ≤ PSR. Related metrics like path length [27] assess similar aspects, however, the tool-call order is additionally taken into account to minimize the duration an object remains in the gripper, thereby reducing the risk of collisions and gripper failures.

Table 1: Generic evaluation model: Task, their generalization and manipulation family

Task	**Generalized task**	**Manipulation family**
Locate the green block	*direct verb*, *direct object*	Specific, direct and single object
Pick-up the green triangle		
Put the red cube on top of the blue cube	*direct verb*, *direct object*, *direct spatial relation*, *direct object*	Specific, direct and multiple objects
Place the cube with the color of the ocean on the paper with the sum of 3 and 4	*direct verb*, *indirect object*, *direct spatial relation*, *indirect object*	Specific, indirect and multi object
Sort the blue and red cube onto the according paper	*indirect verb*, *direct object*, *direct spatial relation*, *indirect object*	
Build a 3-block tower in the order of the Kenya flag	*indirect verb*, *indirect object*, *indirect spatial relation*	
Sort the cubes by color onto the papers	*indirect verb*, *indirect object*, *direct spatial relation*, *indirect object*, *unspecific*	Unspecific, indirect and multi object
Put the red cube on top of the blue cube	*direct verb*, *direct object*, *direct spatial relation*, *direct object*, *unspecific*	

3.3.2 Tasks

To evaluate the LLM agent, a diverse task set is required. Therefore, a generic evaluation framework in which tasks are constructed from fundamental building blocks (Table 1) is proposed. The manipulation taxonomy distinguishes tasks along three dimensions: *specificity* (specific vs. unspecific), *directness* (direct vs. indirect), and *object multiplicity* (single vs. multi-object). This modular design allows researchers to assemble tasks suited to their setups using predefined components. Indirect and unspecific instructions are considered as they are unavoidable in human–robot interaction, highlighting the importance of the more challenging task types at the bottom of Table 1. For task construction, seven elements are defined: direct verb, direct object, direct spatial relation, indirect verb, indirect object, indirect spatial relation, and unspecific. Each task contains at least a verb and an object; spatial relations appear when multiple objects are involved. Direct tasks provide explicit instructions, whereas indirect tasks require inference or simplification (e.g., “cube with the colour of the ocean” → “blue block”). The final two examples in Table 1 illustrate unspecific tasks where environmental context is essential. In task 7, missing information about object colours or quantities prevents direct execution, as the agent would otherwise guess. In task 8, the red cube is unreachable.

Table 2: Summary of the different scenarios and the respective prompt design

Scenario	**Solution examples**	**Multimodal input**	**Environment perception**
I	None	-	Object detection module
II	Manually created text examples	-	Object detection module
III	Auto-generated text examples from unannotated video frames	Video frames	Object detection module

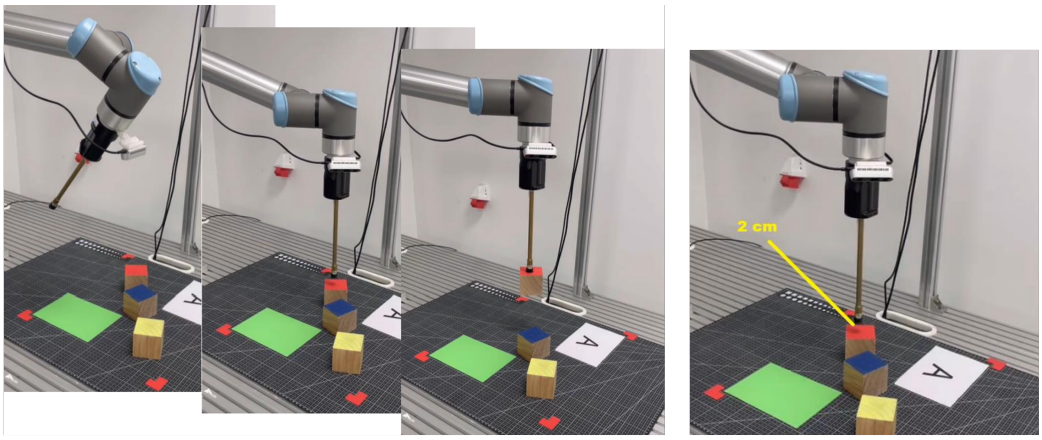
IV	Auto-generated text-examples from annotated video frames	Annotated video frames	Object detection module
V	Manually created text examples	Initial environment image	Object detection module and image captioning

3.3.3 Scenarios

Five performance comparison scenarios are introduced, each with a different prompt for the LLM agent. The scenarios were chosen to explore LLM performance. A zero-shot baseline is established and compared to textual few-shot prompting, as well as three different forms of multimodal few-shot prompting. Table 2 summarizes the different scenarios and the respective prompt designs.

- **Scenario 1** sets a zero-shot baseline where no examples are provided. The LLM agent operates without prior examples, relying solely on the object detection module.
- **Scenario II** evaluates performance using manually created text examples. These examples are provided in pseudo-code format with reasoning steps, as this format performs better than plain text or *json*. Similarly to scenario I, only the object detection module is used.
- **Scenario III** introduces multimodal input in the form of video frames. The LLM auto-generates solution examples based on unannotated video frames of a robot solving a task. Figure 2a illustrates an example sequence of unannotated video frames used in this scenario. The perception of the environment remains based on the object detection module.
- **Scenario IV** builds on scenario III by incorporating annotated video frames. The annotations provide key details, such as correct offsets for grasping objects, as shown in Figure 2b. The LLM uses these annotated frames to generate solution examples, again relying on the object detection module.
- **Scenario V** demonstrates the impact of integrating an image captioning module. This scenario combines the manually created text examples from scenario II with an image captioning module that describes the initial environment, such as available objects. The agent, in this case, has access to both the object detection module and the image captioning system.

In all five scenarios, the agent receives the same *tools.json* file and operates under the same system role.



a) Exemplary frame sequence for picking up the red cube

b) Example annotation of a video frame explaining a 2 cm offset

Figure 2: Visualization of the pick-up process (a) and (b) an annotation highlighting a 2 cm offset

4. Experimental Results

Due to the inherent randomness of LLM-based decision making, the agent's results varied even with the same environment and prompt. Each task was performed 10 times per scenario. The experiments were carried out in August 2024. The results for each metric per respective scenario are shown in Table 3, where

"1.0" indicates success in all 10 cases, and "0.0" indicates complete failure. Among the different scenarios, a brief overview is provided, which will be analyzed in detail in section 4.1.

- **Scenario V** demonstrates the highest overall performance across all three metrics. It particularly outperforms other scenarios in ASR and OPR, demonstrating the most reliable execution, especially for unspecific tasks.
- **Scenarios II** and **scenario IV** rank second, performing well in simpler tasks but struggling with more complex tasks at the end of the table.
- **Scenarios III** and **scenario I** perform similarly badly, with **scenario III** being slightly better than **scenario I**. The weaker performance can be seen particularly in ASR and OPR, where **scenario III** and **scenario I** only perform well for the simplest tasks of locating a single block.

Across all scenarios, several recurring issues were observed: hallucinated plans with unnecessary or missing steps, offset errors in vertical positioning during grasps, stacking failures caused by neglecting object heights, delayed object detection that forces repeated camera movements, and a general lack of environmental awareness in unspecific tasks where additional context is required but not inferred.

Table 3: Experiment Results for the metrics PSR, ASR and OPR for respective scenarios I-V

Task	Plan Success Rate (PSR)					Action Success Rate (ASR)					Optimal Plan Rate (OPR)				
Scenario	I	II	III	IV	V	I	II	III	IV	V	I	II	III	IV	V
Locate the green block.	0.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.6	1.0	1.0	1.0	1.0
Pick-up the green triangle.	0.2	1.0	1.0	1.0	1.0	0.1	1.0	0.0	1.0	1.0	0.2	1.0	1.0	1.0	1.0
Put the red cube on top of the blue cube.	0.8	1.0	1.0	1.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0	1.0	0.0	1.0	1.0
Place the cube with the colour of the ocean on the paper with the sum of 3 and 4.	0.6	1.0	1.0	1.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0	1.0	0.0	1.0	1.0
Sort the blue and red cube onto the according paper.	0.7	1.0	1.0	0.9	1.0	0.0	1.0	0.0	1.0	1.0	0.0	0.8	0.0	0.7	0.9
Build a 3-block tower in the order of the Kenya flag.	0.7	0.8	0.9	0.9	0.9	0.0	0.8	0.0	1.0	1.0	0.0	0.8	0.0	0.8	0.9
Sort the cubes by colour onto the papers.	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.9
Put the red cube on top of the blue cube.	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0

4.1 Findings

Scenario I: The zero-shot agent generates plausible task plans but lacks reliability and parameter accuracy. It achieves a PSR > 0 for 6 of 8 tasks. However, PSR ranges from 0 to 1, indicating high variance in plan generation. Parameter selection is frequently incorrect, particularly for z-coordinates during grasping and stacking, resulting in an ASR of 0 for 6 tasks. Task plans are suboptimal, as objects are not fully localized at initialization and existing structures are not reused, leading to unnecessary tool calls. Unspecific tasks fail due to missing environment information.

Scenario II: Manually defined text examples substantially improve performance. All simple tasks are solved correctly and optimally (ASR = 1). In complex long-horizon tasks, the number of errors decreases significantly, although failures such as hallucinations, delayed error detection, and incorrect stacking still occur. These errors are attributed to long-horizon dependencies, where available information is not consistently applied. While the agent learns effective parameter patterns, it exhibits limited spatial

understanding and mainly reproduces few-shot examples. Unspecific tasks remain unsolved (PSR = 0 for tasks 7 and 8).

Scenario III: Scenario III performs better than the zero-shot baseline but worse than Scenario II. Reliability improves, hallucinations are reduced, and task plans are more coherent, reflected in higher PSR values. For example, tower-building plans correctly reuse existing blocks. However, generated examples lack critical parameter details, such as accurate z-coordinates and initial object localization, resulting in offset and detection failures. Consequently, ASR remains comparable to Scenario I. Unspecific tasks again fail due to insufficient environment information.

Scenario IV: Providing annotated frames yields performance comparable to Scenario II. Differences between both scenarios are minor and attributed to LLM randomness rather than systematic effects. Visual annotations improve performance over Scenarios I and III, particularly in ASR and OPR. However, limitations observed in Scenario II persist, including overfitting to provided examples and reduced robustness in long-horizon tasks. Unspecific tasks still fail (PSR = 0 for tasks 7 and 8).

Scenario V: Scenario V performs on par with Scenarios II and IV for specific tasks but shows a clear improvement for unspecific tasks. Image captioning supplies essential environmental information, resulting in a PSR of 1 for tasks 7 and 8. This scenario demonstrates that multimodal perception is necessary to handle underspecified tasks and cannot be compensated by prompting strategies alone.

4.1.1 Limitations

Experimental findings are limited to the setup used, affected by factors like model choice, temperature setting, and *tools.json* configuration. Testing was done with simple cubes in obstacle-free environments, which may reduce their applicability to more complex settings. The agent relies heavily on the few-shot examples provided and might overfit, which restricts its ability to adapt and generalize across new, untested tasks. Creating and annotating effective examples manually is time-consuming and lacks standardization, making the process dependent on manual effort and expertise. Another limitation is the fact that the agent does not have a true understanding, and instead just successfully copies patterns from examples.

4.1.2 Outlook

To further improve the LLM-integrated robotic system, several experiments and architectural enhancements are proposed. A retrieval-augmented generation database with task-specific examples could reduce reliance on generic responses, mitigate overfitting, and improve adaptability. Limitations in 2D spatial reasoning may be addressed through refined prompting, multimodal inputs such as visual cues with coordinate systems, or the integration of simulation and digital twins that provide explicit spatial references and enable adaptive corrections. For complex tasks, mechanisms to detect and resolve contradictory user instructions are needed, particularly in human–robot interaction, potentially extending the introduced manipulation taxonomy. Although few-shot learning reduces hallucinations, model unpredictability remains a safety concern, requiring constraints and safety checks to ensure reliable operation. Given the rapid progress in VLMs and reasoning models, further improvements are expected in handling ambiguous task descriptions, spatial understanding, and long-horizon planning, including through VLM-enabled video-based learning.

5. Conclusion

This work explores the integration of LLMs into pick-and-place robots, focusing on the effectiveness of zero-shot and multimodal few-shot prompting. Key findings highlight performance improvement of few-shot prompting with multimodal inputs, such as annotated video frames. Image captioning helps robots to be grounded in their environment and solve unspecific tasks. Few-shot prompting faces challenges like trial and error and overfitting. Spatial reasoning remains a key issue for LLM agents. Nevertheless, LLMs enable human-robot interaction through language capabilities and multimodal input processing. This highlights the

potential for future production systems, deploying and adapting robots primarily through natural language instructions and extensive prior knowledge, reducing the need for task-specific programming.

Acknowledgements

This research and development project is funded by the Federal Ministry of Research, Technology and Space (BMFTR) within the "SME-innovative: The Future of Value Creation" funding measure (funding number 02K24K021) and managed by the Project Management Agency Karlsruhe (PTKA). It is also funded by the National Key R&D Program of China, under grant no. 2022YFE0114100. The authors are responsible for the content of this publication.

References

[1] Zhao, W.X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., et al., 2024. A survey of large language models. arXiv preprint arXiv:2303.18223.

[2] Sahoo, P., Singh, A.K., Saha, S., Jain, V., Mondal, S., Chadha, A., 2024. A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv preprint arXiv:2402.07927.

[3] Zhang, C., Chen, J., Li, J., Peng, Y., Mao, Z., 2023. Large language models for human–robot interaction: A review. Biomimetic Intelligence and Robotics 3 (4), 100131.

[4] Liu, H., Zhu, Y., Kato, K., Tsukahara, A., Kondo, I., Aoyama, T., et al., 2024. Enhancing the LLM-based robot manipulation through human-robot collaboration. IEEE Robotics and Automation Letters 9, 6904–6911.

[5] Wolber, J., Muyang, L., Koch, D., Bretz, L., Lanza, G., Baer, F., 2025. Large Language Models for Robotics: A systematic literature review on prompt engineering. In: Min, J., Zhang, W., Fleischer, J., Lanza, G. (Eds.), Sustainable Manufacturing Innovations: Focus on New Energy Vehicles, Production Robots, and Software-Defined Manufacturing. Lecture Notes in Production Engineering, ICSM 2024. Springer, Cham.

[6] Shanahan, M., 2023. Talking about large language models. arXiv preprint arXiv:2212.03551.

[7] Amaratunga, T., 2023. Understanding Large Language Models: Learning Their Underlying Concepts and Technologies. Apress, Berkeley, CA.

[8] Aljanabi, M., Yaseen, M.G., Ali, A.H., Mohammed, M.A., 2023. Prompt engineering: Guiding the way to effective large language models. Iraqi Journal for Computer Science and Mathematics, 151–155.

[9] Kojima, T., Gu, S.S., Reid, M., Matsuo, Y., Iwasawa, Y., 2023. Large language models are zero-shot reasoners. arXiv preprint arXiv:2205.11916.

[10] Schick, T., Dwivedi-Yu, J., Dessi, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., et al., 2023. Toolformer: Language models can teach themselves to use tools. arXiv preprint arXiv:2302.04761.

[11] Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., et al., 2024. A survey on large language model based autonomous agents. Frontiers of Computer Science 18 (6).

[12] Huang, D., Yan, C., Li, Q., Peng, X., 2024. From large language models to large multimodal models: A literature review. Applied Sciences 14 (12).

[13] Zhang, J., Huang, J., Jin, S., Lu, S., 2024. Vision-language models for vision tasks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 46 (8), 5625–5644.

[14] Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., et al., 2021. Learning transferable visual models from natural language supervision. arXiv preprint arXiv:2103.00020.

[15] Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., et al., 2024. Grounding DINO: Marrying DINO with grounded pre-training for open-set object detection. arXiv preprint arXiv:2303.05499.

[16] Karli, U.B., Chen, J.-T., Antony, V.N., Huang, C.-M., 2024. Alchemist: LLM-aided end-user development of robot applications. Association for Computing Machinery, New York.

[17] Singh, I., Blukis, V., Mousavian, A., Goyal, A., Xu, D., Tremblay, J., et al., 2022. ProgPrompt: Generating situated robot task plans using large language models. arXiv preprint arXiv:2209.11302.

[18] Jin, Y., Li, D., A, Y., Shi, J., Hao, P., Sun, F., et al., 2023. RobotGPT: Robot manipulation learning from ChatGPT. arXiv preprint arXiv:2312.01421.

[19] Vemprala, S., Bonatti, R., Bucker, A., Kapoor, A., 2023. ChatGPT for robotics: Design principles and model abilities. arXiv preprint arXiv:2306.17582.

[20] Wake, N., Kanehira, A., Sasabuchi, K., Takamatsu, J., Ikeuchi, K., 2023. ChatGPT empowered long-step robot control in various environments: A case application. IEEE Access 11, 95060–95078.

[21] Wang, B., Zhang, J., Dong, S., Fang, I., Feng, C., 2024. VLM see, robot do: Human demo video to robot action plan via vision-language model. arXiv preprint arXiv:2410.08792.

[22] Ahn, M., Brohan, A., Brown, N., Chebotar, Y., Cortes, O., David, B., et al., 2022. Do as I can, not as I say: Grounding language in robotic affordances. arXiv preprint arXiv:2204.01691.

[23] Lin, K., Agia, C., Migimatsu, T., Pavone, M., Bogh, J., 2023. Text2Motion: From natural language instructions to feasible plans. Autonomous Robots, 1345–1365.

[24] Huang, W., Xia, F., Xiao, T., Chan, H., Liang, J., Florence, P., et al., 2022. Inner Monologue: Embodied reasoning through planning with language models. arXiv preprint arXiv:2207.05608.

[25] Bhat, V., Kaypak, A.U., Krishnamurthy, P., Karri, R., Khorrami, F., 2024. Grounding LLMs for robot task planning using closed-loop state feedback. arXiv preprint arXiv:2402.08546.

[26] Luan, Z., Lai, Y., Huang, R., Bai, S., Zhang, Y., Zhang, H., et al., 2024. Enhancing robot task planning and execution through multi-layer large language models. Sensors 24 (5).

[27] Yang, Z., Ning, L., Wang, H., Jiang, T., Zhang, S., Cui, S., et al., 2024. Text2Reaction: Enabling reactive task planning using large language models. IEEE Robotics and Automation Letters 9 (5), 4003–4010.

[28] Wang, C., Hasler, S., Tanneberg, D., Ocker, F., Joublin, F., Ceravola, A., et al., 2024. LaMI: Large language models for multi-modal human-robot interaction. In: CHI Conference on Human Factors in Computing Systems Extended Abstracts. ACM, New York.

[29] Mei, A., Zhu, G.-N., Zhang, H., Gan, Z., 2024. ReplanVLM: Replanning robotic tasks with visual language models. IEEE Robotics and Automation Letters 9 (11), 10201–10208.

[30] Kannan, S.S., Venkatesh, V.L.N., Min, B.-C., 2024. SMART-LLM: Smart multi-agent robot task planning using large language models. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 12140–12147.

[31] Huang, W., Wang, C., Zhang, R., Li, Y., Wu, J., Fei-Fei, L., 2023. VoxPoser: Composable 3D value maps for robotic manipulation with language models. arXiv preprint arXiv:2307.05973.

[32] Xia, W., Wang, D., Pang, X., Wang, Z., Zhao, B., Hu, D., et al., 2024. Kinematic-aware prompting for generalizable articulated object manipulation with LLMs. arXiv preprint arXiv:2311.02847.

[33] Shentu, Y., Wu, P., Rajeswaran, A., Abbeel, P., 2024. From LLMs to actions: Latent codes as bridges in hierarchical robot control. arXiv preprint arXiv:2405.0479.

[34] Yang, J., Zhang, H., Li, F., Zou, X., Li, C., Gao, J., 2023. Set-of-Mark Prompting unleashes extraordinary visual grounding in GPT-4V. arXiv preprint arXiv:2310.11441.

[35] Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Chen, X., Choromanski, K., et al., 2023. RT-2: Vision-language-action models transfer web knowledge to robotic control. arXiv preprint arXiv:2307.15818.

[36] Song, C.H., Wu, J., Washington, C., Sadler, B.M., Chao, W.-L., Su, Y., 2023. LLM-Planner: Few-shot grounded planning for embodied agents with large language models. arXiv preprint arXiv:2212.04088.

[37] Sarch, G., Wu, Y., Tarr, M.J., Fragkiadaki, K., 2023. Open-ended instructable embodied agents with memory-augmented large language models. arXiv preprint arXiv:2310.15127.

[38] Zhou, Z., Song, J., Yao, K., Shu, Z., Ma, L., 2023. ISR-LLM: Iterative self-refined large language model for long-horizon sequential task planning. arXiv preprint arXiv:2308.13724.

[39] Sakib, M.S., Sun, Y., 2024. Consolidating trees of robotic plans generated using large language models to improve reliability. International Journal of Artificial Intelligence and Robotics Research 1 (1), 2450002.

[40] Zhang, Y., Wang, Z., Zhang, S., Peng, Y., Chen, M., 2023. Boosting robot intelligence in practice: Enhancing robot task planning with large language models. In: International Conference on Robotics and Automation Engineering (ICRAE), 90–94.

[41] Liu, B., Jiang, Y., Zhang, X., Liu, Q., Zhang, S., Biswas, J., Stone, P., 2023. LLM+P: Empowering large language models with optimal planning proficiency. arXiv preprint arXiv:2304.11477.

[42] Bai, J., Bai, S., Chu, Y., Cui, Z., Dang, K., Deng, X., et al., 2023. Qwen technical report. arXiv preprint arXiv:2309.16609.

[43] OpenAI, 2024. GPT-4o.

[44] Aryan, S., Subrahmanya, S., Jagtap, P., Mani, L., R., S.M., 2023. Image captioner: Captioning for visual impact. In: International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE), 1–8.

[45] Zhao, C., Yuan, S., Jiang, C., Cai, J., Yu, H., Wang, M.Y., et al., 2023. ERRA: An embodied representation and reasoning architecture for long-horizon language-conditioned manipulation tasks. arXiv preprint arXiv:2304.02251.

[46] Rana, K., Haviland, J., Garg, S., Abou-Chakra, J., Reid, I., Suenderhauf, N., 2023. SayPlan: Grounding large language models using 3D scene graphs for scalable robot task planning. arXiv preprint arXiv:2307.06135.

[47] Chen, L., Liu, Z., He, W., Li, Y., Luo, R., Yang, M., 2024. Long context is not long at all: A prospector of long-dependency data for large language models. arXiv preprint arXiv:2405.17915.

[48] Chen, Y., Pesaranghader, A., Sadhu, T., Yi, D.H., 2024. Can we rely on LLM agents to draft long-horizon plans? Let's take TravelPlanner as an example. arXiv preprint arXiv:2408.06318.

Biography

Dominik Koch (*1996) received his B.Sc. in Mechanical Engineering and his M.Sc. in Mechatronics and Information Technology from Karlsruhe Institute of Technology. He is currently a research associate at the wbk Institute of Production Science, pursuing his doctoral degree. His work focuses on robot-based optical inspection in circular manufacturing.

Jakob Wolber (*1999) is working as a Consultant at the Boston Consulting Group in Munich and is part of BCG's Operations practice. He holds a bachelor's and a master's degree in industrial engineering from the Karlsruhe Institute of Technology (KIT), with a focus on production engineering.

Zhou Shi (*1997) has been applied AI researcher at Siemens Ltd., China since 2022. Zhuo's work mainly focuses on researching and applying cutting edge AI technology in the automation area, including computer vision solution for pick & place and anomaly detection use cases, Large Language Model for domain knowledge retrieval, Vision Language Action model for dexterous robot control.

Ji Bo Cheng (*1995) has been researcher at Siemens Ltd., China since 2023. Bo Cheng's work mainly focuses on researching and transforming cutting edge technology into real scenarios, including computer vision solution for pick & place, whole body control of humanoids robot. His research interest consists of kinematic, dynamic of industry robot, generate AI and VLA.

Lucas Bretz (*1992) is General Manager of the Global Advanced Manufacturing Institute (GAMI) in Suzhou, China. He holds a PhD in Mechanical Engineering from the Karlsruhe Institute of Technology. His work focuses on operational excellence, digital transformation, and AI driven production systems. He develops robotics and automation applications for factory optimization and integrates large language models into secure industrial environments.

Alexander Geiser (*1994) received his M.Sc. in Mechatronics from HS Mannheim in 2022. Previously, he worked as a Technical Project Leader in special machine construction. Currently, he is a research associate at the wbk Institute of Production Science, pursuing his doctoral degree (Dr.-Ing.).

Felix Baer (*1992) has been the Head of the Autonomous Factory Department at Siemens Ltd., China since 2022. He has conducted research in advanced robotics for 14 years and holds several patents in this field. As a PhD candidate, he focuses on business transformation driven by generative AI products in industry.

Martin Benfer (*1994) is Deputy Head of Production Systems at wbk Institute of Production Science at Karlsruhe Institute of Technology. He leads the Production Systems Planning and Quality Assurance groups. His research focuses on data and model-based decision-making, digital twins, and resilience in production systems.

Florian Stamer (*1991) has been Professor of Production Management at Leuphana University Lüneburg since 2025. Before that, he led a junior research group and served as chief engineer in quality assurance at Karlsruhe Institute of Technology. His research links AI with production and quality management across the product life cycle. He focuses on digital production systems, production networks, and data-driven decision support for sustainable and efficient manufacturing

Gisela Lanza (*1970) is Professor of Production Systems and a member of the board of directors at the wbk Institute of Production Science at Karlsruhe Institute of Technology. She leads the Production Systems division, focusing on global production strategies, production system planning, and quality assurance. Her expertise centers on data-driven design and evaluation of production systems, ensuring robust and efficient manufacturing operations.

8th Conference on Production Systems and Logistics

A Structured Literature Review and Interview study of human-centred Challenges in Artificial Intelligence Implementation

Verena Dondorf[1], Kitrujan Kokilanathan[1], Johanna P. Schenk[1], Daniela Vogel[1], Roman Dumitrescu[2]

[1]*Fraunhofer Institute for Mechatronic Systems Design, Paderborn, Germany*
[2]*Chair for Advanced Systems Engineering, Heinz Nixdorf Institute, Paderborn, Germany*

Abstract

Artificial intelligence offers substantial performance and innovation potential, yet many initiatives struggle because human centred challenges create organisational resistance. This paper investigates two research questions that link problem diagnosis and solution design in the context of AI implementation and change management, drawing on qualitative interviews with fourteen participants and a systematic literature review of fifty-six studies. The first research question addresses the problem statement by identifying and structuring key clusters of human-centred challenges associated with AI initiatives in organisations. The results outline four overarching challenges and related clusters. Exemplary challenges include the central role of leadership in guiding AI related change processes and the impact of demographic differences on perceptions, skills and acceptance of AI technologies. The second research question forms the solution statement. Building on the identified overarching challenges, the study examines which organisational requirements must be fulfilled to adequately address them. In addition, it develops a set of actionable measures and specifies how these measures need to be designed and applied to respond effectively to the challenges outlined in the problem statement and to reduce organisational resistance. The paper concludes with future research, recommending that subsequent studies elaborate and validate the identified requirements from the solution statement in greater detail to further support sustainable AI implementation.

Keywords

AI; human-centred; implementation; challenges; digital transformation; change management

1. Introduction

Artificial Intelligence (AI) is demonstrated as a pivotal future technology [1] that continues to shape various industries [2]. Its application in business contexts provides a competitive advantage by enabling more informed and sustainable decision-making [3] and unlocking new revenue opportunities [4,5].

Consequently, companies are compelled to engage with this disruptive technology to harness its potential [6]. Supporting this view, a study by KPMG reveals that 83% of worldwide companies are committed to increasing their investments in AI [7].

Based on the above assumptions, one would expect the use of AI in companies to continue increasing steadily. However, according to a survey conducted by the German Federal Statistical Office, only one in five companies in Germany currently makes use of AI technologies [8]. Moreover, the definition of AI usage remains unclear in this context. The AI Index Report indicates that, according to respondents, 78% of companies reported using AI in 2024 that is an increase of 23% within one year [2]. Nevertheless, global uptake is still short of full saturation. In addition, most surveys record AI deployment in only one

DOI: https://doi.org/10.15488/20981
ISSN: 2701-6277

organisational area, suggesting that many business functions have yet to benefit from AI solutions. Furthermore, recent labor market evidence indicates that current AI adoption in professional tasks still falls far short of what would be technically feasible, so that substantial productivity and innovation potentials remain unused [8].

One central explanation for the delayed adoption may lie in the complexity of AI implementation. In line with this, recent work on AI in industrial work argues that in contrast to earlier digital technologies, AI systems increasingly operate as semi-autonomous actors in socio-technical work systems. This amplifies uncertainty and raises new demands on human-AI collaboration and work design [6]. The literature describes numerous challenges associated with AI integration, as exemplified by BRANDTER. The main challenges identified include issues with data integration and quality, organisational resistance to change, a shortage of skilled personnel, as well as concerns regarding the opacity of AI systems and the reliability of their outputs [9]. SHANG further highlights the costs of implementation and maintenance, the importance of top-down support, and the competencies of employees as additional critical factors [10]. AI implementation is therefore complex not only due to the technological intricacies involved, which were primarily focused on in the past, but also because employees need to be adequately supported, as AI often evokes uncertainty and fear among staff [11]. A BCG summary of the 2024 World Economic Forum in Davos underscores that AI represents a business challenge rather than a purely technical one, raising questions such as how to motivate employees to embrace AI [12]. Consequently, this paper focuses exclusively on human-centred challenges to examine them in greater depth.

Building on this focus, it is imperative to acknowledge that human-centred challenges frequently impede the progression of AI projects. Consequently, it is crucial to ascertain the factors that are of significance to employees. The introduction of AI requires consideration not only of the technological component but also of the human factor. According to SMITH, the biggest challenge in AI adoption may lie in the acceptance of the technology by people [13]. Empirical studies show a strong link between resistance to organisational change and the willingness to adopt AI [14]. Furthermore, ethical concerns such as privacy, bias and discrimination, safety, and security can hinder the AI implementation [15]. Decision-makers, who are responsible for introducing AI (e.g. management or project managers), must therefore identify the causes of resistance and define targeted strategies to overcome them. These causes can be diverse, as BANTA explains in his study, noting that individuals who do not continuously improve their IT skills or lack experience in the field are at greater risk and tend to face more difficulties with AI implementation [16]. If employees are not adequately engaged during the rollout, fear may rise, and acceptance of AI can drop. In extreme cases, the entire change initiative may stall or even be abandoned.

To avoid the failure of an AI initiative due to human-centred challenges, companies must address these issues and implement suitable change initiatives. To introduce AI, companies must understand the challenges relating to AI and people and know how to respond to them to reap the efficiency benefits of AI following successful implementation. It is evident that some companies remain unaware of or overwhelmed by these challenges, consequently hindering their capacity to respond effectively. For example, trust in AI, as well as a well-founded understanding of the generated outputs, constitutes a fundamental prerequisite for acceptance and use of the system [17]. In addition, user-friendly and comprehensible explanations contribute substantially to fostering the cognitive process of understanding when working with AI systems [18]. Moreover, many individuals report that they are afraid of losing their job due to AI [19].

Thus, a structured and strategically aligned change management (CM) approach is suitable for supporting a targeted AI implementation by systematically addressing both: people and the required processes while taking the associated challenges into account. The implementation of AI has the potential to contribute to the establishment of an innovative-oriented culture, the advancement of leadership capabilities, and the enhancement of trust in AI systems [20–22]. The most important aspect of this endeavour is the identification of the most pertinent points for the company and the subsequent development of a response to them. Currently, there are countless change measures intended to provide support, but these do not guarantee that the relevant challenges will be eliminated. Evidence from the literature suggests that suitable qualification and professional developments measures are often lacking, despite being essential for strengthening

employees' AI-related knowledge. Accordingly, it is necessary to align these measures with AI-specific requirements [23]. There are various ways in which organisations can use measures to respond to the challenges. In CM, the dissemination of information, participation, and the qualification of employees are frequently cited as success factors, as confirmed by various authors [24–27]. It is therefore imperative for companies to utilise these diverse methods correctly during the implementation of AI to facilitate communication, training and participation.

The literature has prompted the following two research questions (RQ). The first provides an overview of the overwhelming challenges related to employees and the associated clustering of AI. The second RQ posits a series of requirements that will be of particular significance for companies in the future, as they seek to ascertain the status quo of the challenges they face. The RQ also sets out specific ways to effect change in response to the challenges:

1.1 What human-centred challenges of AI implementation and change management in organisations lead to organisational resistance?

1.2 Which requirements and actionable measures arise from these challenges for AI implementation?

2. Methodology

To address the two RQ, this paper adopts a mixed-methods approach comprising a systematic literature review (SLR) and interviews. The following Figure 1 outlines the characteristics of both methods and the procedures employed.

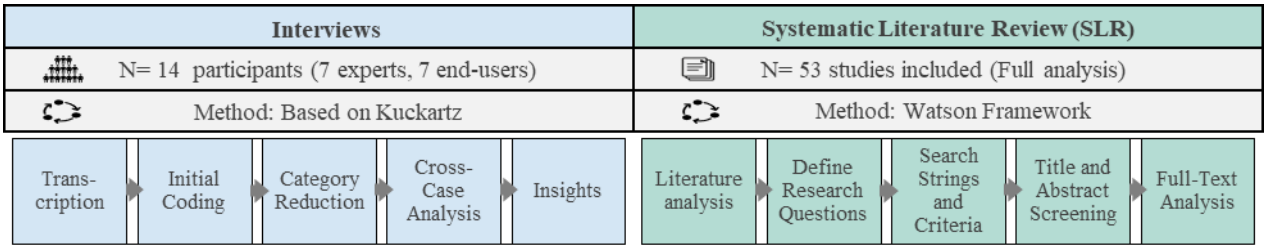


Figure 1: Procedure of Interviews and SLR

At the outset, two RQ were derived from gaps in the existing literature, forming the foundation of the review design. The SLR procedures are delineated first, followed by the procedures for the interviews.

Following WATSON's **SLR** protocol as a guiding framework, an iterative search strategy was applied in Scopus. An initial test with a minimal keyword set (e.g., "Artificial Intelligence" AND "Implementation") yielded predominantly generic results, prompting the addition of further domain-specific terms [20]. After several refinement cycles, five search strings were finalised, each combining the following keywords in different permutations in English and German:

- Artificial Intelligence or AI
- Generative Artificial Intelligence or GenAI
- Implementation or Introduction or Change
- Employee or Human
- Influencing Factors or Challenges or Success Factors
- Maturity Model or Readiness Model or Readiness Check
- Quantitative Assessment

These searches targeted peer-reviewed articles, conference papers, and books published between 2020 and 2025 and returned 747 records (see Figure 2). After duplicate removal, 641 titles were screened independently by two researchers. During title and abstract screening, numerous studies focused on education or healthcare. Due to their limited relevance for corporate AI implementation, they were excluded. 53 papers met all inclusion criteria and were available and were selected for full-text analysis, with 19 additional pertinent publications from a partner university added to the final dataset as a manual search result.

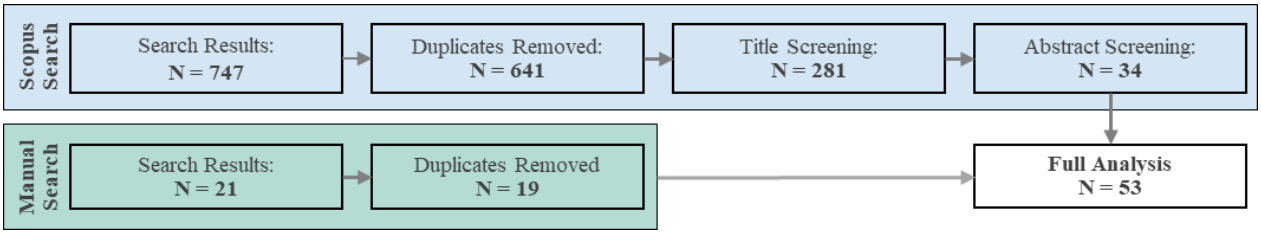


Figure 2: Number of relevant Papers identified

Guided by the RQ, the relevant articles were examined through a structured data extraction process. The primary aim was to identify the relevant challenges with a strong focus on the human perspective. In this paper, 'challenges' are defined as tasks for the company. For example, if there is a lack of trust in AI, the company's challenge is to 'build trust'. Whenever it was unclear whether a challenge had truly addressed people, this issue was discussed in relation to the interview results. All pertinent findings were entered into a standardised extraction sheet and consolidated into a comparative matrix.

The **interview** procedure followed the principles of KUCKARTZ's qualitative content analysis [21]. Participants were recruited from an expert group and an end-user group to enable a triangulate integration of perspectives. The expert group comprised individuals with several years of professional experience in the field of AI, while the end-user group included participants from diverse industries who reported a general openness toward AI. In total, N = 14 participants were interviewed, with 7 participants in each group, using online, in-person, and written format.

Data were collected using guideline-based, semi-structured interviews with an intended duration of 30 minutes. The interview guide covered questions on AI implementation, including comparisons between AI and classical IT systems, and incorporated a timeline exercise to elicit when participants perceived the highest need for support across the implementation process. In addition, interviewees were asked to classify proposed measures into information, qualification, and participation formats and to assess their intensity before, during, and after implementation using three levels. Intensity is defined as the priority assigned to the change format, thereby ensuring that the execution addresses the challenges.

Information formats refer to measures that provide continuous updates on the status of the change process (e.g., newsletters, FAQ). Qualification formats comprise capability-building measures intended to strengthen employees' AI-related knowledge and skills (e.g., workshops, trainings) and participation formats capture opportunities for active involvement and co-creation (e.g., idea management and feedback mechanisms).

3. Results and Conceptual Framework

The results of the analysis are presented in this chapter. The results of the SLR and the interviews are presented together in one overview and explained in the text.

It opens with descriptive statistics of the SLR, including an overview of how many papers were published in each year (see Table 1). The SLR was conducted between April and June 2025. It is evident that the topic is gaining increasing relevance, as evidenced by the proliferation of publications in this field.

Year of publication					
2020	2021	2022	2023	2024	2025*
4	4	6	8	24	7

Table 1: Number of publications per year included in the SLR *until May 2025

3.1 Visualization of the Problem Statement of Core Challenges in AI Adoption

Firstly, the results of the primary RQ are listed: ***What human-centred challenges of AI implementation and CM in organisations lead to organisational resistance?*** (Figure 3). The figure enumerates all the challenges that were mentioned in the interviews and the SLR. These were divided into two categories: overarching challenges and content-related challenges. The term 'overarching' is used in a non-specific manner,

signifying that these challenges do not have a substantive focus. Instead, they are overarching and provide an orientation on how to respond to the content-related challenges.

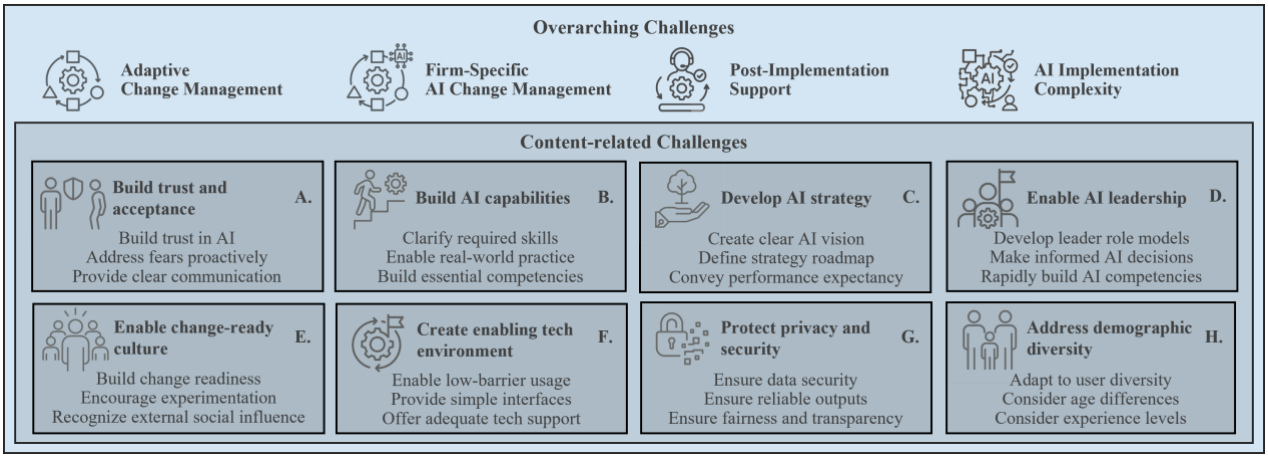


Figure 3: Overview of overarching and content-related challenges in AI implementation

It is evident from the interview evidence that four overarching human-centred challenges should orient AI implementation summarised in Figure 3. **Firstly**, it is imperative that CM is characterised by dynamism, iteration, transparency, and adaption. During the interviews, emphasis was placed on the fact that the considerable variability inherent in AI systems renders comprehensive planning for a rollout at the outset impracticable. The impracticality of meticulous planning in the field of AI necessitates the adoption of agile methodologies, characterised by frequent learning cycles, transparent communication, and the capacity for dynamic adjustments. **Secondly**, it is imperative that CM be tailored to AI and the specific organisation. As system autonomy increases, employees require more information and reassurance. It is important to note that each organisation is unique in terms of its level of engagement with new technologies. It is evident that a 'one size fits all' approach is not efficacious in the context of AI. **Thirdly**, it is imperative that support extends beyond the initial deployment phase. To function effectively, organisations require consistent and readily identifiable points of contact, ongoing training programmes, and effective mechanisms to manage model updates and policy changes following the implementation phase. **Fourthly**, the implementation of AI is characterised by its complexity, which links technology, processes, skills, governance, ethics, and trust. Cross-Industry evidence from the World Economic Forum (WEF) similarly shows that the main obstacles to realizing AI‘s potential are organizational rather than purely technical, stressing the need for new operating models, leadership practices, skills and trust mechanisms when embedding AI into core workflows [22]. Survey data from German industrial companies underlines this and indicate that employees perceive the introduction of AI as more uncertain, more complex and more demanding in terms of data security than the implementation of other digital technologies [23].

It is vital for organisations to prioritise these four overarching challenges when introducing AI. The interviews also revealed several challenges related to the content. The subsequent section is a compilation of these items, which are then integrated with the findings from the SLR into eight clusters (Table 2).

Code	Explanation	Sources
A.	To facilitate the effective integration of AI into the workplace, it is imperative that employees possess a high degree of **trust and acceptance**. SHARMA reports that 25% of respondents fear job loss, which demonstrates the importance of clear communication and visible safeguards [19]. During the interviews, communication was identified as "the most significant factor" in the implementation of AI. This finding suggests that transparent, continuous communication is pivotal in reducing uncertainty.	[19,24–37]
B.	To achieve a more profound comprehension and utilisation of AI, it is imperative that individuals possess the necessary **competencies**. Organisations require role-based competency maps, targeted training, and guided practice on actual tasks.	[10,16,27,32,33,36–45]

C.	A clear **Strategy** gives direction and meaning. Employees should understand why AI is introduced and what results are expected. WIJAYA emphasises that performance expectancy must be clearly defined, ensuring that all parties have a clear understanding of the benefits AI can bring to the workplace [46].	[10,33,36–38,40,42,45–48]
D.	**Leaders** must act as role models, make informed decisions, and build their own literacy quickly. ALHUSBAN demonstrates that leaders frequently determine AI decisions without prior contextual knowledge [43], which may result in suboptimal outcomes.	[10,11,32,33,35–37,43,49–51]
E.	**Organisational readiness** encompasses culture and social dynamics. Teams need explicit permission, sufficient time, and appropriate incentives to experiment and learn. Empirical evidence indicates that an innovation-oriented culture amplifies the impact of AI use [52].	[10,19,25,27,31,33,36,37,39,41,45,50,52]
F.	A supportive **technology environment** removes scepticism. The challenge is to provide simple interfaces, low-barrier access, and seamless integration. The interviews confirmed that from a technological perspective, usability barriers in work, including prompting and the competent handling of AI tools, create friction that can translate into frustration.	[11,26,35,39,45,49,53]
G.	**Privacy, security, and ethics** are a core challenge for AI adoption. Across domains, automated decisions create risks of unequal treatment, opacity, and data exposure. Recruiting is one illustrative high-stakes case: when AI participates in screening or selection, studies report concerns about disparate impact on protected groups [54].	[27,33–36,38,40,42,43,45,47,50,54,55]
H.	The adoption process is influenced by disparities in **demographic characteristics**. KIM finds that perceptions vary by age and professional experience, which affects confidence and learning needs [11]. Consequently, rollout plans should be differentiated.	[11,41,47,56]

Table 2: Content-related Challenges

The results of the study demonstrate the specific challenges that companies encounter when introducing AI in relation to employees, as outlined in the problem statement.

3.2 Extended Framework Integrating Organizational Solution Statement

The results for the second RQ are presented in the following chapter: ***Which requirements and actionable measures arise from these challenges for AI implementation?*** (Figure 4). Building on the challenges identified for RQ1, the figure derives practical support for companies in two forms: general approaches (Requirements for the Challenges) and actionable measures. Each item is grounded in evidence from the SLR and the interviews and linked to the challenges.

Responding to the interview results and the overarching challenges in Figure 4, the next section proposes a Solution Statement for organisations planning and operating AI. The Solution Statement focuses on four requirements that directly address adaptive CM, firm specific CM, the need for post implementation support, and the high complexity of AI. The requirements provide a clear overview of how the company can respond to the overarching challenges posed by RQ1. **First**, organisations need a structured readiness assessment to establish the status quo. Employees must be able to voice their views in their own words, and the instrument should provide filter options such as age to reveal differences and enable tailored actions. It cannot be assumed that employees will openly communicate their concerns, such as fears of job loss. Therefore, external assessments, for example by managers, may fail to accurately reflect the true status quo. **Second**, the assessment must be repeatable at regular intervals and simple to run so that shifts in sentiment become visible and the organisation can respond quickly. **Third**, companies require an overview of the measures that can help produce concrete action recommendations so that decision makers know how to respond throughout the process and which measures remain useful after full implementation. **Fourth**, the assessment must absorb the complexity of AI by covering the full set of human related factors, ensuring that trust, skills, roles, communication, governance, and ethics are considered in an integrated way. Existing readiness assessments evaluate a range of factors related to AI implementation. However, this approach is insufficient to capture the full complexity of human behaviour, as it does not account for all relevant dimensions. As a result, the limited number of questions could lead to an incomplete assessment.

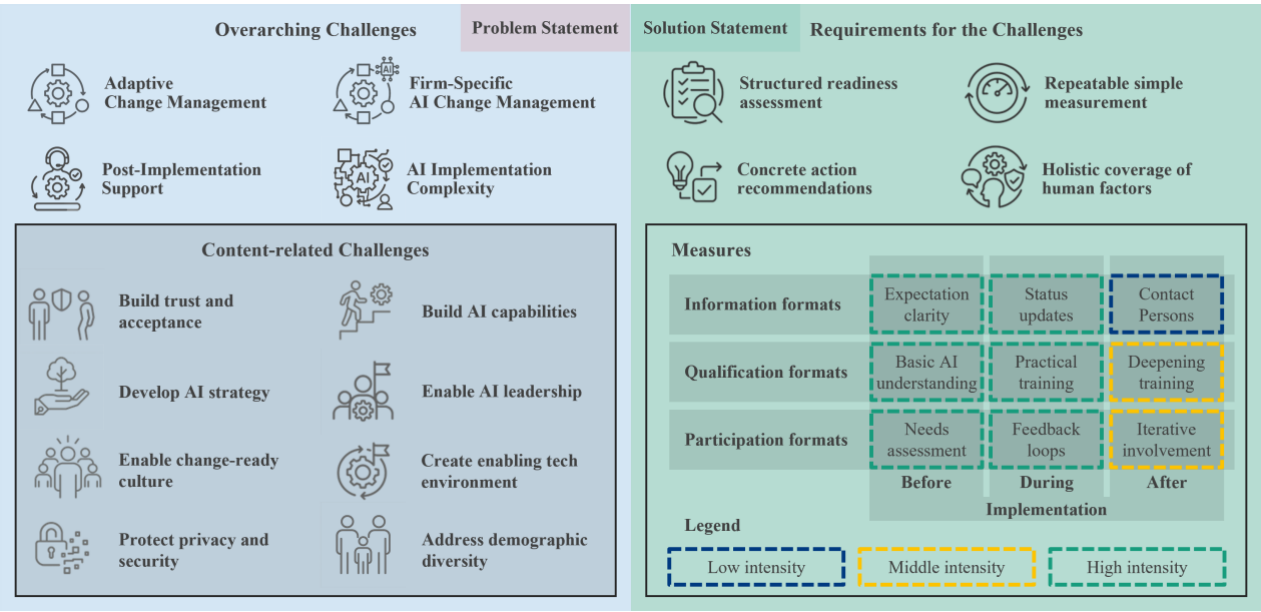


Figure 4: Overview of AI implementation challenges and corresponding solution approaches

The measures in Figure 4 summarise which change formats are useful across the three implementation phases and their recommended intensity, linking each to RQ1 challenges. Regarding **information formats**, the pre-implementation focus is on expectation clarity to reduce uncertainty and resistance, and to set realistic AI expectations. This is achieved through clear, continuous communication that makes the AI strategy and vision explicit and aligns leadership. During the implementation phase, regular status updates ensure transparency and communicate safeguards for privacy and security to build trust. Following implementation, accessible contact people resolve daily questions and maintain coherent communication, stabilising long-term acceptance. **Qualification formats** directly build AI capabilities. Before implementation, foundational training addresses skill gaps. During implementation, task-focused practice and safe testing with sandbox and tooling access support an enabling tech ecosystem and create implementation readiness. After implementation, follow-up training consolidates knowledge, offers advanced paths, and remains accessible to diverse groups, strengthening responsible use. Regarding **participation formats**, pre-implementation needs assessments and use anonymous channels to surface fears and feedback. During the implementation phase, feedback loops facilitate continuous adjustment, ensuring the system's adaptability to evolving requirements. Following implementation, iterative involvement translates experience into improvements that reinforce a change-ready culture and sustain trust.

4. Discussion and future research

Drawing on interviews and a SLR, this study shows why many organisations struggle with the introduction of AI and how human-centred challenges and CM translate into resistance. These findings also clarify why the adoption of AI often fails. Organisations must address multiple challenges simultaneously, yet few do so comprehensively because these challenges cannot be solved quickly. Leadership gaps require sustained capacity building through multi-year development, coaching and practice. Structural realignment and team norms also cannot be rushed. External forces also influence adoption. Social influence extends beyond the workplace, and employees may carry anxieties about job loss from the media, their peers, and their professional networks. Therefore, firms must address both internal capabilities and the broader social context. There is no one-size-fits-all blueprint. Effective approaches must be tailored to organisational contingencies, such as industry and openness for AI. Qualitative and survey evidence from engineering contexts underscores this need for tailoring by showing that successful AI introduction depends on continuous stakeholder participation, a shared terminology and stepwise, realistic implementation rather than top-down technology rollouts [57].

For the second RQ the study develops a solution statement that translates the challenges into requirements and actionable measures. It shows that organisations must rethink AI implementation by moving away from long-term detailed planning toward a flexible and human-centred approach, because with AI new challenges such as fear and scepticism strongly influence adoption. The four identified requirements of this shift by enabling organisations to systematically assess how strongly specific challenges are expressed among employees rather than relying on managerial assumptions. This highlights that AI rollouts cannot be fully specified in advance but require an iterative approach with repeated assessments over time. Such assessments make it possible to adapt measures as challenges evolve and to keep the focus on human related factors such as trust, skills, and roles.

The findings further indicate that the measures identified must have very high priority throughout AI implementation and require sustained commitment to changing initiatives. A key tension is resourcing, since many organisations underinvest in change because it does not directly increase technical efficiency, even though it enables adoption, correct use, trust, and business value. Future research should validate the proposed requirements and measures set through longitudinal and comparative studies and combine quantitative indicators with employee narratives to offer concrete guidance for organisations. Future studies should focus particularly on how such an assessment should be carried out, what else needs to be considered, and what the final assessment should look like. This is important for the successful implementation of AI.

In addition, the change measures will result in further research in the future. For now, this paper can only discuss them and conduct an initial evaluation, which must be considered. Future studies should examine the measures in greater depth and provide a clear evaluation of their impact on the challenges.

The novelty of this research is underscored by the recency of the reviewed studies, most published in 2023 or later. The concentration of new work signals the urgency of the topic. Successful AI implementation requires equal attention to technology and to people.

4.1 Limitation

This study offers valuable insights. However, there are a few limitations. First, the analysis relies solely on Scopus, which may omit relevant work indexed elsewhere. Future research should integrate EBSCOhost, Web of Science, and Google Scholar to build a more comprehensive evidence base. Second, the interview sample consisted mainly of technology enthusiasts, which introduces selection bias and limits representativeness. Future studies should recruit participants from a broader range of roles, sectors, and levels of technical affinity, including more critically minded respondents. Third, challenges are difficult to disentangle across perspectives. The study focused only on employee-related challenges, yet in practice technological constraints can shape employee experiences and behaviours. This interdependence should be examined in follow-up work using qualitative or quantitative designs.

5. Conclusion

In conclusion, this study shows that AI introduction often falters because human-centred challenges are numerous and mutually reinforcing. The study explored with Interviews and SLR two RQ. RQ1: What human-centred challenges of AI implementation and CM in organizations lead to organisational resistance? The findings map four overarching challenges together with eight challenge clusters, which explain where resistance arises and why organisations can quickly become overwhelmed. RQ2: Which requirements and actionable measures arise from these challenges for AI implementation? Organisations need procedures and tools to diagnose their specific challenge profile and should consistently apply change formats focused on information, participation and qualification to support adoption. Future work should refine and test these instruments with companies and employees through quantitative studies to enable scaled AI.

Note: Artificial intelligence was used to enhance linguistic processing. All content has been carefully checked by the authors.

References

[1] Streim, A., Czachowski, L., 2025. KI-Nutzung boomt – aber die Angst vor Abhängigkeit vom Ausland ist groß. https://www.bitkom.org/Presse/Presseinformation/KI-Nutzung-boomt-Angst-vor-Abhaengigkeit-Ausland-gross#_. Accessed 3 April 2026.

[2] Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R., Clark, J., 2024. Artificial Intelligence Index Report 2024. https://hai.stanford.edu/assets/files/hai_ai-index-report-2024-smaller2.pdf. Accessed 3 April 2026.

[3] Ali, M., Khan, T.I., Khattak, M.N., ŞENER, İ., 2024. Synergizing AI and business: Maximizing innovation, creativity, decision precision, and operational efficiency in high-tech enterprises. Journal of Open Innovation: Technology, Market, and Complexity 10 (3), 100352.

[4] Daugherty, P.R., Wilson, H.J., Narain, K., 2023. Generative AI Will Enhance — Not Erase — Customer Service Jobs. https://hbr.org/2023/03/generative-ai-will-enhance-not-erase-customer-service-jobs. Accessed 16 June 2025.

[5] Clauss, T., 2017. Measuring business model innovation: conceptualization, scale development, and proof of performance. R & D Management 47 (3), 385–403.

[6] Gabriel, S., Falkowski, T., Graunke, J., Dumitrescu, R., Murrenhoff, A., Kretzschmer, V., Hompel, M.t., 2024. Künstliche Intelligenz und industrielle Arbeit - Perspektiven und Gestaltungsoptionen.

[7] Höck, B., Klein, L., Madan, A., 2025. Generative KI in der deutschen Wirtschaft 2025: Wo stehen deutsche Unternehmen bei der Implementierung generativer KI? https://kpmg.com/de/de/home/themen/2025/04/studie-generative-ki-in-der-deutschen-wirtschaft-2025.html.

[8] Massenkoff, M., McCrory, P., 2026. Labor market impacts of AI: A new measure and early evidence. https://www.anthropic.com/research/labor-market-impacts. Accessed 31 March 2026.

[9] Brandtner, P., 2024. Digital Transformation of Supply Chain Management - Challenges and Strategies for Successfully Implementing Data Analytics in Practice, in: Proceedings of the 2024 8th International Conference on E-Commerce, E-Business, and E-Government. ICEEG 2024: 2024 the 8th International Conference on E-Commerce, E-Business, and E-Government, Ajman United Arab Emirates. 28 05 2024 30 05 2024. ACM, New York, NY, USA, pp. 36–42.

[10] Shang, G., Low, S.P., Lim, X.Y.V., 2023. Prospects, drivers of and barriers to artificial intelligence adoption in project management. BEPAM 13 (5), 629–645.

[11] Kim, Y., Blazquez, V., Oh, T., 2024. Determinants of Generative AI System Adoption and Usage Behavior in Korean Companies: Applying the UTAUT Model. Behavioral sciences (Basel, Switzerland) 14 (11).

[12] Schweizer, C., 2025. The Four Signals That Broke Through at Davos.

[13] Smith, M., Prabhakar, G., Nisar, T.M., Tseng, H.-T., 2025. Adoption of AI by the HR function in the civil service. ER 47 (1), 104–126.

[14] Frick, N.R.J., Mirbabaie, M., Stieglitz, S., Salomon, J., 2021. Maneuvering through the stormy seas of digital transformation: the impact of empowering leadership on the AI readiness of enterprises. Journal of Decision Systems 30 (2-3), 235–258.

[15] Stahl, B.C., Brooks, L., Hatzakis, T., Santiago, N., Wright, D., 2023. Exploring ethics and human rights in artificial intelligence – A Delphi study. Technological Forecasting and Social Change 191, 122502.

[16] Banța, V.-C., Rîndașu, S.-M., Tănasie, A., Cojocaru, D., 2022. Artificial Intelligence in the Accounting of International Busi-nesses: A Perception-Based Approach. Sustainability 14 (11), 6632.

[17] Holzinger, A., 2021. The Next Frontier: AI We Can Really Trust, in: Kamp, M., Koprinska, I., Bibal, A., Bouadi, T., Frénay, B., Galárraga, L., Oramas, J., Adilova, L., Krishnamurthy, Y., Kang, B., Largeron, C., Lijffijt, J., Viard, T., Welke, P., Ruocco, M., Aune, E., Gallicchio, C., Schiele, G., Pernkopf, F., Blott, M., Fröning, H., Schindler, G., Guidotti, R., Monreale, A., Rinzivillo, S., Biecek, P., Ntoutsi, E., Pechenizkiy, M., Rosenhahn, B., Buckley, C., Cialfi, D., Lanillos, P., Ramstead, M., Verbelen, T., Ferreira, P.M., Andresini, G., Malerba, D.,

Medeiros, I., Fournier-Viger, P., Nawaz, M.S., Ventura, S., Sun, M., Zhou, M., Bitetta, V., Bordino, I., Ferretti, A., Gullo, F., Ponti, G., Severini, L., Ribeiro, R., Gama, J., Gavaldà, R., Cooper, L., Ghazaleh, N., Richiardi, J., Roqueiro, D., Saldana Miranda, D., Sechidis, K., Graça, G. (Eds.), Machine Learning and Principles and Practice of Knowledge Discovery in Databases, vol. 1524. Springer International Publishing, Cham, pp. 427–440.

[18] Dodge, J., Liao, Q.V., Zhang, Y., Bellamy, R.K.E., Dugan, C., 2019. Explaining models, in: Proceedings of the 24th International Conference on Intelligent User Interfaces. IUI '19: 24th International Conference on Intelligent User Interfaces, Marina del Ray California. 17 03 2019 20 03 2019. ACM, New York, NY, USA, pp. 275–285.

[19] Sharma, R., Reddy, B., 2024. Strategic Approaches to AI Integration in Organizations: Managing Change and Maximizing Benefits, in: 2024 4th Asian Conference on Innovation in Technology (ASIANCON). 2024 4th Asian Conference on Innovation in Technology (ASIANCON), Pimari Chinchwad, India. 23.08.2024 - 25.08.2024. IEEE, pp. 1–5.

[20] Xiao, Y., Watson, M., 2019. Guidance on Conducting a Systematic Literature Review. Journal of Planning Education and Research 39 (1), 93–112.

[21] Kuckartz, U., Rädiker, S., 2024. Fokussierte Interviewanalyse mit MAXQDA. Springer Fachmedien Wiesbaden, Wiesbaden.

[22] Novo, F.G., Hoang, J., O´Regan, K., 2026. Organizational Transformation in the Age of AI: How Organizations Maximize AI's Potential. https://www.weforum.org/publications/organizational-transformation-in-the-age-of-ai-how-organizations-maximize-ais-potential/. Accessed 31 March 2026.

[23] Dondorf, V., Lebedeva, E., Thommes, K., Dumitrescu, R., 2025. Evolution von KI in der industriellen Arbeitswelt. Fraunhofer IEM.

[24] López-Solís, O., Luzuriaga-Jaramillo, A., Bedoya-Jara, M., Naranjo-Santamaría, J., Bonilla-Jurado, D., Acosta-Vargas, P., 2025. Effect of Generative Artificial Intelligence on Strategic Decision-Making in Entrepreneurial Business Initiatives: A Systematic Literature Review. Administrative Sciences 15 (2), 66.

[25] Hussain, M.A., Gupta, R., Kushwaha, A., Samanta, P., Khulbe, M., Ahmad, V., 2024. Transforming Technology for Online Marketing with Focus on Artificial Intelligence: A Qualitative Approach, in: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India. 24.06.2024 - 28.06.2024. IEEE, pp. 1–5.

[26] Abel, D.H., Rusilowati, U., Firmansyah, F., Dian Wulandari, O.A., Lindzani, N.A., Astuti, E.D., 2024. Comparative Analysis of Technological Integration in Human Resources Management During Industry 4.0, in: 2024 3rd International Conference on Creative Communication and Innovative Technology (ICCIT). 2024 3rd International Conference on Creative Communication and Innovative Technology (ICCIT), Tangerang, Indonesia. 07.08.2024 - 08.08.2024. IEEE, pp. 1–6.

[27] Hua, D., Petrina, N., Young, N., Cho, J.-G., Poon, S.K., 2024. Understanding the factors influencing acceptability of AI in medical imaging domains among healthcare professionals: A scoping review. Artificial intelligence in medicine 147, 102698.

[28] Kumar, S., Mittal, S., 2024. Employee learning and skilling in AI embedded organizations – predictors and outcomes. DLO 38 (4), 11–15.

[29] Sedkaoui, S., Benaichouba, R., 2024. Generative AI as a transformative force for innovation: a review of opportunities, applications and challenges. EJIM.

[30] Gupta, N., Gaur, V., Bhatt, A., Gaur, S., Srivastava, N., 2024. Intelligent Techniques for Managing Employee Performance: Addressing Psychosocial Challenges, in: 2024 International Conference on Intelligent & Innovative Practices in Engineering & Management (IIPEM). 2024 International Conference on Intelligent & Innovative Practices in Engineering & Management (IIPEM), Singapore, Singapore. 25.11.2024 - 25.11.2024. IEEE, pp. 1–6.

[31] Suseno, Y., Chang, C., Hudik, M., Fang, E.S., 2022. Beliefs, anxiety and change readiness for artificial intelligence adoption among human resource managers: the moderating role of high-performance work systems. The International Journal of Human Resource Management 33 (6), 1209–1236.

[32] Nair, M., Svedberg, P., Larsson, I., Nygren, J.M., 2024. A comprehensive overview of barriers and strategies for AI implementation in healthcare: Mixed-method design. PloS one 19 (8), e0305949.

[33] Nortje, M.A., Grobbelaar, S.S., 2020. A Framework for the Implementation of Artificial Intelligence in Business Enterprises: A Readiness Model, in: 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, United Kingdom. 15.06.2020 - 17.06.2020. IEEE, pp. 1–10.

[34] Ridho, W.F., 2023. An Examination of the Opportunities and Challenges of Conversational Artificial Intelligence in Small and Medium Enterprises. RoBES 11 (3), 6–17.

[35] Siradhana, N.K., Arora, R.G., 2024. Examining the Influence of Artificial Intelligence Implementation in HRM Practices Using T-O-E Model. Vision: The Journal of Business Perspective.

[36] Jöhnk, J., Weißert, M., Wyrtki, K., 2021. Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors. Bus Inf Syst Eng 63 (1), 5–20.

[37] Armutat, S., Wattenberg, M., Mauritz, N., 2024. Designing an Artificial Intelligence Maturity Model for Human Resources (HR-AIMM). ICAIR 4 (1), 50–58.

[38] Alshehhi, E.T., Khadem, M., 2024. Investigating the AI Drivers Integrating in Healthcare Systems, in: 2024 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD). 2024 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Sharjah, United Arab Emirates. 04.11.2024 - 06.11.2024. IEEE, pp. 1–6.

[39] Khatri, S., Pandey, D.K., Penkar, D., Ramani, J., 2020. Impact of Artificial Intelligence on Human Resources, in: Sharma, N., Chakrabarti, A., Balas, V.E. (Eds.), Data Management, Analytics and Innovation, vol. 1016. Springer Singapore, Singapore, pp. 365–376.

[40] Lugito, N.P.H., Cucunawangsih, C., Suryadinata, N., Kurniawan, A., Wijayanto, R., Sungono, V., Sabran, M.Z., Albert, N., Budianto, C.J., Rubismo, K.Y., Purushotama, N.B.S.A., Zebua, A., 2024. Readiness, knowledge, and perception towards artificial intelligence of medical students at faculty of medicine, Pelita Harapan University, Indonesia: a cross sectional study. BMC medical education 24 (1), 1044.

[41] Barodi, M., Lalaoui, S., 2025. The Readiness of Civil Servants to Join the Era of Artificial Intelligence: A Case Study of Moroccan Public Administration. Change Management: An International Journal 25 (1), 1–21.

[42] Calvo, J., 2025. Towards an Integrative Framework for AI Readiness and Market Disruption: The AI-Readiness and Market Dynamics Disruption Index (AIM-DDI) for Competitive AI Leadership, in: Taratukhin, V., Levchenko, A., Kim, S. (Eds.), Cultural Perspectives of Human-Centered and Technological Innovations, vol. 2233. Springer Nature Switzerland, Cham, pp. 135–154.

[43] Alhusban, M.I., Khatatbeh, I.N., Alshurafat, H., 2025. Exploring the influence, implications and challenges of integrating generative artificial intelligence into organizational learning and development. Competitiveness Review An International Business Journal (2), 21–24.

[44] Ahn, M.J., Chen, Y.-C., 2022. Digital transformation toward AI-augmented public administration: The perception of government employees and the willingness to use AI in government. Government Information Quarterly 39 (2), 101664.

[45] Limat, C., 2022. Disruptionspotenzial künstlicher Intelligenz: Ein Reifegradmodell zur Einführung ganzheitlicher KI-Initiativen in Unternehmen. Wirtsch Inform Manag 14 (1), 60–67.

[46] Wijaya, T.T., Su, M., Cao, Y., Weinhandl, R., Houghton, T., 2025. Examining Chinese preservice mathematics teachers' adoption of AI chatbots for learning: Unpacking perspectives through the UTAUT2 model. Educ Inf Technol 30 (2), 1387–1415.

[47] Ramadan, O.M.E., Alruwaili, M.M., Alruwaili, A.N., Elsehrawy, M.G., Alanazi, S., 2024. Facilitators and barriers to AI adoption in nursing practice: a qualitative study of registered nurses' perspectives. BMC nursing 23 (1), 891.

[48] Dolle, N., Kuzmina-Merlino, I., 2023. Possibilities and Barriers in Implementing of AI-Based Automation Techniques in Management, in: Kabashkin, I., Yatskiv, I., Prentkovskis, O. (Eds.), Reliability and Statistics in Transportation and Communication, vol. 640. Springer International Publishing, Cham, pp. 414–423.

[49] Najdawi, A., 2020. Assessing AI Readiness Across Organizations: The Case of UAE, in: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India. 01.07.2020 - 03.07.2020. IEEE, pp. 1–5.

[50] Marcinkevage, C., Kumar, A., 2025. Generative AI in Higher Education Constituent Relationship Management (CRM): Opportunities, Challenges, and Implementation Strategies. Computers 14 (3), 101.

[51] Vărzaru, A.A., 2022. An Empirical Framework for Assessing the Digital Technologies Users' Acceptance in Project Management. Electronics 11 (23), 3872.

[52] Zhou, J., 2024. The influence of artificial intelligence technology application on employee work performance: Based on the intermediary role of enterprise innovation level, in: Proceedings of the 2024 7th International Conference on Information Management and Management Science. IMMS 2024: 2024 the 7th International Conference on Information Management and Management Science, Beijing China. 23 08 2024 25 08 2024. ACM, New York, NY, USA, pp. 281–286.

[53] Russo, D., 2024. Navigating the Complexity of Generative AI Adoption in Software Engineering. ACM Trans. Softw. Eng. Methodol. 33 (5), 1–50.

[54] Mihai, F., Aleca, O.E., Gheorghe, M., 2023. Digital Transformation Based on AI Technologies in European Union Organizations. Electronics 12 (11), 2386.

[55] Yun, C.H., Teoh, A.P., Khaw, T.Y., 2024. Artificial Intelligence Integration in e-Government: Insights from the Korean Case, in: 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA). 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China. 27.02.2024 - 29.02.2024. IEEE, pp. 1159–1164.

[56] MARINAȘ, L.E., PĂUN, C.V., DIACONESCU, M., SMIRNA, T.G., 2024. Artificial Intelligence Readiness and Employment: A Global Panel Analysis. ECECSR 58 (4/2024), 57–74.

[57] Humpert, L., Tihlarik, A., Wäschle, M., Anacker, H., Dumitrescu, R., Albers, A., Albert, B., Röbenack, S., Pfeiffer, S., 2023. Investigating the potential of artificial intelligence for the employee from the perspective of AI-experts, in: 2023 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD). 2023 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Rabat, Morocco. 22.11.2023 - 24.11.2023. IEEE, pp. 1–8.

Verena Dondorf studied business psychology in Bielefeld. Afterward, she worked for several years as a team leader at a logistics company and is currently a research associate at Fraunhofer IEM, focusing on change management and digital transformation.

Kitrujan Kokilanathan studied industrial engineering with a focus on mechanical engineering at the Technical University of Berlin. He is currently a research associate at the Fraunhofer Institute for Mechatronic Systems Design IEM, focusing on technology and innovation management.

Johanna P. Schenk completed a bachelor's degree in business psychology and a master's degree in management information systems. Her areas of focus include digital transformation, change management, artificial intelligence, and technology acceptance. Professionally, she has been working for several years as a lecturer at a university of applied sciences.

Dr. Daniela Vogel studied industrial engineering with a focus on mechanical engineering at the University of Paderborn. She currently heads the Strategy and Transformation Management group at the Fraunhofer Institute for Mechatronic Systems Design IEM. There she leads research and industry projects in the fields of New Work, Smart Factory, digitalization and sustainability strategy.

Prof. Dr.-Ing. Roman Dumitrescu is Director at the Fraunhofer Institute for Mechatronic Design Technology IEM and Head of the Advanced Systems Engineering department at Paderborn University. His research focuses on the product development of intelligent technical systems. Prof. Dumitrescu is also Managing Director of the technology network Intelligent Technical Systems OstWestfalenLippe (it's OWL). Among other things, he is a member of the expert group of the German government's Innovation Dialogue, the research advisory board of the 3-D MID e. V. research association, the steering committee of the NRW state government's Economy & Work 4.0 initiative and the administrative board of the RKW Competence Center.

8th Conference on Production Systems and Logistics

Wear Volume Measurement Of Cutting Tools: Structured Literature Review

Benjamin Thorenz[1], Johannes Rauch[1], Frank Döpper[1,2]

[1] *University of Bayreuth, Chair of Manufacturing and Remanufacturing Technology, Bayreuth, Germany*
[2] *Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Bayreuth, Germany*

Abstract

Machining processes on cutting machine tools, which according to DIN 8580 can be classified under the main manufacturing group of cutting, specifically cutting with a geometrically defined cutting edge, are used in almost every manufacturing company. The tools are subject to unavoidable wear during use. Wear patterns occur on the cutting edge, which depend on the materials of the working pair and the machining conditions.

Recognizing this progressive wear on tools is important because it significantly affects manufacturing quality and cost-effectiveness. To avoid unplanned downtime and quality losses, tools are often replaced after predefined usage cycles in industrial practice. These often conservatively chosen replacement intervals mean that tools are replaced too early, without exploiting the full potential of their entire tool life.

One theory is that companies choose this approach because they cannot adequately quantify tool wear, which could be attributed to the classic wear parameters (1D) used in industrial practice, such as width of flank wear land. One-dimensional measurement methods cannot accurately represent the three-dimensional wear of tools, as they do not take into account factors like the spatial extent and wear volume. Three-dimensional measurement techniques overcome these limitations by generating detailed 3D data of the cutting edges. For this reason, this paper presents the findings of a structured literature review on the topic of wear volume measurement of cutting tools. These show that most previous publications merely describe wear volume. An interpretation of wear volume in comparison to classic wear parameters appears only rudimentary. The formation of correlations with wear effects was lacking in earlier publications, which is why further research is necessary.

Keywords

Wear measurement; wear volume measurement; width of flank wear land; cutting tool; structured literature review

1. Introduction

Cutting tools are subject to unavoidable wear during use. Wear patterns occur on the cutting edge, which depend on the materials of the working pair and the machining conditions. The active parts of cutting tools – the cutting edges – are subjected to *mechanical*, *thermal*, and *chemical stresses* during the machining process, which cause wear on the tool cutting edges. These stresses can be stationary, i.e., constant over time (turning in smooth cutting or drilling), or transient, i.e., varying over time (milling or turning in interrupted cutting). [1]

In industrial practice, cutting tools are often replaced after predefined usage cycles to avoid unplanned downtime and quality losses. Since these replacement intervals are typically chosen conservatively, cutting

DOI: https://doi.org/10.15488/20982
ISSN: 2701-6277

tools are frequently exchanged too early, without fully utilizing their potential tool life. This makes recognizing the progressive wear of the tool essential, as it directly affects manufacturing quality and cost-effectiveness. [2, 3, 4, 5]

A possible explanation is that the approach described is chosen because tool wear cannot be adequately quantified. This may result from the traditional wear parameters used in industrial practice, such as the width of flank wear land, or from the loss of overview regarding the development and interactions of the individual deterioration phenomena due to their diverse combinations. Wear volume measurement could provide a remedy. To this end, the methodology for volumetric quantification of tool wear must be capable of taking into account both the complexity of the tool geometry and the advanced state of wear of the tool [6].

After introducing the topic and highlighting the research question, the theoretical background to tool wear is presented (Section 2). This is followed by the research approach in Section 3, which consists of a structured literature review used to derive a concept matrix. In Section 4, the most important results are discussed. The paper concludes with a summary and outlook to motivate future research approaches in Section 5.

2. State of the art – wear measurement

Depending on the type and duration of the load, the wear patterns that occur on the cutting part during the machining process vary in severity [2]. There are a variety of classic wear patterns that need to be interpreted and measured, such as flank wear and crater wear [7]. In addition, the diverse combination of these classic wear parameters can lead to a loss of overview of the development and interactions between the individual deterioration phenomena.

Flank wear is mainly caused by mechanical friction at the edge of the major and minor cutting edges. It leaves a visible mark with a width VB with vertical wear lines, see Figure 1a). This results in a displacement of the cutting edge SV, which is the amount by which a tool must be readjusted if it is to achieve the same level of wear as an unused cutting edge, see Figure 1b).

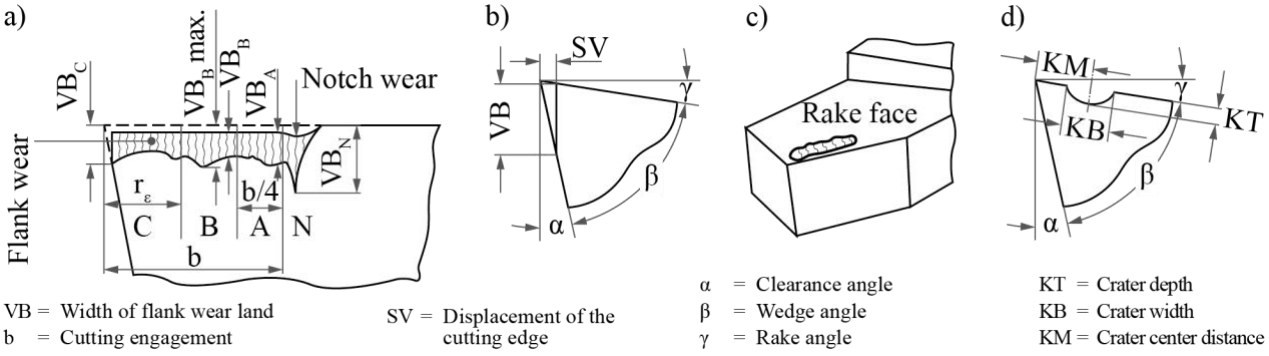


Figure 1: Flank wear (a, b) and crater wear (c, d), according to [8, 9]

The crater wear shown in Figure 1c) and d) appears as a trough-shaped hollow on the rake face. It is caused by the interaction of friction and diffusion and changes the chip flow direction.

Another criticism of classic wear parameters, such as width of flank wear land, which is commonly used to interpret tool condition, is the varying quantification of tool wear. [6] Figure 2 serves as an example of a contradictory scenario. The uneven friction patterns on the flank face in Figure 2a) and Figure 2b), as well as the part worn by material removal in Figure 2c), are all quantified with the same VB value. Such inconsistencies can also be easily demonstrated quantitatively.

3D measurement techniques enable the generation of three-dimensional data sets of the cutting edges. By superimposing the data sets (cutting edge in new condition/unworn and after the machining tests (worn) or at defined steps in between), the material loss can be quantified, providing a measure of wear volume [10].

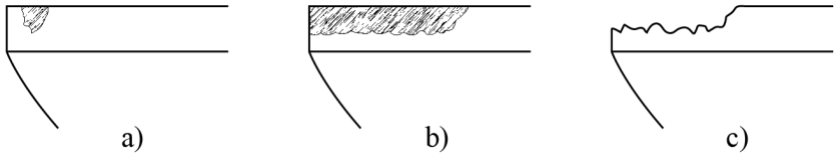


Figure 2: An inconsistent wear quantification scenario [6]

3. Analyzing the literature

3.1 Structured literature review

The literature review followed the methodology proposed by WEBSTER & WATSON [11]. Accordingly, they recommend a three-part approach to identifying the source material for review [11]:

1. Search leading journals and selected conference proceedings.
2. Consider previous articles by searching backwards through the citations from step 1.
3. Use a forward search to identify articles that reference the articles identified in the previous steps.

According to COOPER's taxonomy [12], the review aims to provide research findings for general scientists by identifying key topics within a representative literature review.

The literature search was conducted systematically using databases and search media such as Web of Science, Google Scholar, Elicit, and Litmaps, as well as subject-specific databases of the University Library of Bayreuth. Primary search terms covering wear volume and measurement-related aspects were systematically combined with process-specific supplementary search terms and with supplementary search terms addressing comparisons between one-dimensional and three-dimensional wear metrics using the Boolean operator AND. All primary search terms were combined with all supplementary search terms. This procedure resulted in a total of 154 search combinations. The complete list of search terms is provided in the Appendix. In February 2026, an initial set of 36 publications was found based on the analysis of titles and abstracts. Following a more detailed assessment of the full publication content, 25 articles were selected as relevant. A total of three iterations were performed, with two researchers independently evaluating the literature in each case.

3.2 Concept matrix

Following the suggestions of WEBSTER & WATSON [11] and building on SALIPANTE ET AL. [13], a concept matrix was developed for analyzing and synthesizing the literature. In line with the research question, this consists of four characteristics that provide insights into the state of the literature. What they all have in common is that the sources deal with the measurement of wear volume (VV), the volumetric material loss of the cutting edge compared to the unworn tool, which is why this is not listed separately as a category.

Category 1 asks whether the respective source also deals with the classic wear parameters (1D) described in Section 2 – width of flank wear land (VB) or crater wear (KT, KB). The second category is related to the first and describes whether a comparison is made between classic wear parameters and wear volume – regardless of what this looks like (this category will be specified in more detail later in the paper, see Table 2). Category 3 asks which machining process was used to obtain the findings. Information about the tool, specifically about cutting materials and coatings, is provided in the following category. The last category also asks about the material used, as this has a decisive influence on the development of wear.

Table 1 below provides an overview of the individual studies before they are briefly described and answers the categories listed.

Table 1: Concept matrix for scientific work and approaches to wear volume

Year	Author	1D	Comparison with 3D	Machi-ning process	Tool information		Material
					cutting material	coating	
1999	MOON & LING [14]	✗	✗	Turning	Carbide	N/A	AISI 1040
2004	DAWSON & KURFESS [15]	✗	✗	Turning	PCBN	✓ , ✗	AISI 52100
2006	WANG ET AL. [16]	✗	✗	Milling	Cermet, carbide	✓	N/A
2009	BURGER ET AL. [17]	✗	✓ (VB)	Milling	Carbide	✓	Titanium
2012	KUTTOLAMADOM ET AL. [6]	✓	✓ (VB)	Turning, Milling	Carbide	✓	Ti-6Al-4V
2012	KUTTOLAMADOM ET AL. [18]	✓	✓ (VB)	Milling	Carbide	✓	Ti-6Al-4V
2014	LIU & XIONG [19]	✗	✗	Turning	N/A	N/A	N/A
2015	ČERČE ET AL. [20]	✓	✓ (VB, KT)	Turning	Carbide	✓ (several)	100Cr6
2015	KUTTOLAMADOM ET AL. [21]	✗	✗	Turning	Carbide	✗	Ti-6Al-4V
2015	KUTTOLAMADOM ET AL. [22]	✗	✗	Milling	Carbide	✓	Ti-6Al-4V
2017	SEID AHMED ET AL. [23]	✓	✗	Turning	Carbide	✓ (several), ✗	S32750
2018	BOING ET AL. [24]	✗	✗	Turning	PCBN	✓	AISI 4340, 52100
2018	CASTRO ET AL. [25]	✓	✗	Turning	PCBN	✓	AISI 52100
2019	CAPASSO ET AL. [26]	✓	✓ (VB)	Turning	Carbide	✓ (several)	Inconel DA718
2020	BOING ET AL. [27]	✓	✓ (VB)	Turning	PCBN	✓	AISI 52100
2021	BAI ET AL. [28]	✓	✗	Drilling	Carbide	N/A	CFRP
2022	YU ET AL. [29]	✓	✓ (KB)	Turning	Carbide	✓	GH4202
2023	PENG ET AL. [30]	✗	✗	Milling	N/A	N/A	N/A
2023	WANG ET AL. [31]	✓	✓ (KB)	Turning	PCBN	✓ (several), ✗	GH4202
2024	CHOWDHURY ET AL. [32]	✓	✓ (VB)	Turning	Carbide	✓	Ti6Al4V
2024	MBAKOP NANSHIE ET AL. [33]	✗	✗	N/A	Tool steels	✓	Wood
2024	QIN ET AL. [10]	✓	✓ (VB)	Milling	Carbide	✓	AISI 1045
2024	TIAN ET AL. [34]	✓	✓ (VB)	Milling	HSS	N/A	C45
2025	LI ET AL. [35]	✓	✓ (VB)	Milling	N/A	N/A	N/A
2026	LUKÁŠ ET AL. [36]	✗	✗	Turning	Carbide	✗	Ti6Al4V

- MOON & LING introduced a method for three-dimensional measurement of tool wear based on the shape-from-focus (SFF) technique. Using a microscope and image processing technology, the three-

dimensional shape of the cutting edge was reconstructed from a sequence of images with varying lens distances. The application case was the turning of carbon steel. [14]

- DAWSON ET AL. used white light interferometry to quantify wear volume separately on the rake and flank faces of a CBN-tipped tool. Cutting forces increased with flank wear but remained stable or decreasing under dominant crater wear. [15]
- WANG ET AL. presented a method for non-contact crater wear measurement based on the phase shift method. A high-precision 3D profile of the tool is created by projecting four color stripe patterns onto the cutting surface and capturing and analyzing grayscale images. As a result, it is possible to determine various parameters of crater wear. The system was tested for the milling process. [16]
- To determine the wear volume, BURGER ET AL. from Clemson University compared solid models of worn indexable inserts with those of an unused cutting insert. For this purpose, point clouds were recorded using a white light interferometer and transformed into CAD models. [17]
- In subsequent studies at Clemson University, KUTTOLAMADOM ET AL. qualitatively assessed conventional wear metrics and developed a standardized volumetric method, validated by ANOVA and machining tests on Ti-6Al-4V. [6, 18] The three-dimensional evaluation method was applied to turning tools, investigating the influence of different carbide grain sizes on wear volume [21, 22].
- Another method for measuring crater wear is based on a stereo vision system with a CCD camera and Matlab-based software. LIU & XIONG developed image matching algorithms to estimate the wear volume on the rake face. This makes it possible to determine the volume and crater depth. Rotational tests confirmed the feasibility of this method. [19]
- Using laser profile sensors, ČERČE ET AL. were able to precisely determine three-dimensional wear profiles and thus record tool wear in three dimensions. Based on the measured and calculated wear volume, a model was developed that maps the machining performance. The system was tested in a turning process involving bearing steel. [20]
- SEID AHMED ET AL. investigated and compared the tool life, wear mechanisms, and tribological performance of coated and uncoated carbide indexable inserts when turning super duplex stainless steel. A three-dimensional wear evaluation was performed using focus variation method. [23]
- CASTRO ET AL. and BOING ET AL. proposed a wear rate method based on wear volume measured by focus variation method. The method was tested by both in a turning process for carbon steel, with BOING ET AL. also investigating a low-alloy chromium-nickel-molybdenum steel. [24, 25] BOING ET AL. further defined 3D wear parameters to analyze PCBN tool wear in hardened steels [24] and developed a tool life prediction method, identifying wear volume as a suitable criterion [27].
- CAPASSO ET AL. investigated the wear properties of two different coating systems on carbide tools during the turning of Inconel DA718. The wear behavior was characterized based on the progression of wear volume and flank wear. In addition, cutting forces and wear mechanisms were analyzed. In this context, measurements were performed using focus variation method to compare wear volume and the development of built-up edges. [26]
- Using a semi-analytical model, BAI ET AL. were able to predict the wear behavior when drilling carbon fiber reinforced plastic (CFRP) with carbide drills. Focus variation method was used as the measuring device to generate 3D data. The wear volume was derived from calculated cutting edge profiles. [28]
- YU ET AL. evaluated the influence of micro-textures on cutting inserts during turning of the superalloy GH4202 with respect to tool wear behavior. Focus variation method was employed to quantify wear volume and the volume of adhered material. [29]
- A three-dimensional reconstruction method for milling tool tips was presented by PENG ET AL., using the non-subsampled discrete shearlet transform (NSST) and the Laplace algorithm together with SFF principle to obtain an initial 3D model from image sequences. In a second step, depth information was refined by fitting the pixel focus curve in the image sequence with Gaussian

function. Finally, the volume of the resulting 3D point cloud, and thus the wear volume in the damaged region, was calculated using Simpson's double numerical integration. [30]

- WANG ET AL. investigated the wear mechanisms of coated PCBN tools during high-speed turning of superalloy GH4202, measuring wear volume using focus variation method. [31]
- CHOWDHURY ET AL. investigated wear volume in wet finish turning of Ti-6Al-4V using CrN-coated carbide cutting tools to evaluate the influence of coating deposition parameters on wear mechanisms and tool life. Wear volumes associated with built-up edge and crater wear were measured by focus variation method and referenced to the original tool geometry. [32]
- MBAKOP NANSHIE ET AL. measured wear volume of wood-cutting tools using an optical 3D surface scanning microscope, allowing a direct comparison of wear resistance across different tool steels and PVD coatings. [33]
- QIN ET AL. evaluated the wear condition of end mills based on wear volume at the tool tip during milling of AISI 1045 steel. The wear volume was determined from a 3D point cloud of the tool tip reconstructed using an optical SSF principle, where multi-focus image sequences were processed to obtain depth information and calculate wear volume as a quantitative indicator. [10]
- TIAN ET AL. proposed a three-dimensional wear region reconstruction method for end mills with an in-situ dual-sensor system combining machine vision and laser scanning. The wear region was extracted using a combined threshold segmentation method, and the three-dimensional morphology was reconstructed via a shape-from-shading (SFS) algorithm enhanced by laser-based depth measurements. Experimental validation was carried out in milling tests on C45 carbon steel. [34]
- LI ET AL. presented a milling tool wear detection method using optical microscopy reconstruction to generate fully focused images and depth maps for wear topography. [35]
- LUKÁŠ ET AL. analyzed wear volume during turning of Ti-6Al-4V using focus variation method to assess crater-dominated wear under different tool clearance angles and cutting conditions. [36]

As already mentioned, the comparison of 1D and 3D is further specified, as the individual comparisons showed significant differences. These ranged from simple plots or tabular representations of wear development values (designated "P" in Table 2) to attempts at interpretation ("I"), which described the curve progressions at least qualitatively. In addition, the sources were examined for correlations with wear effects ("C"), such as the increase in cutting force, cutting temperature, changed chip formation, and deterioration of surface quality with increasing wear [37]. Table 2 below summarizes this investigation and further breaks down the comparison of wear parameters carried out in Table 1.

Table 2: Concept matrix augmented with units of analysis for scientific work and approaches, compare 1D and 3D

Year	Author	1D	Comparison with 3D		
			P	I	C
2009	BURGER ET AL. [17]	✗	✓	✗	✗
2012	KUTTOLAMADOM ET AL. [6]	✓	✓	✗	✗
2012	KUTTOLAMADOM ET AL. [18]	✓	✓	✓	✗
2015	ČERČE ET AL. [20]	✓	✓	✗	✗
2019	CAPASSO ET AL. [26]	✓	✓	✗	✗
2020	BOING ET AL. [27]	✓	✓	✓	✗
2022	YU ET AL. [29]	✓	✓	✗	✗
2023	WANG ET AL. [31]	✓	✓	✗	✗
2024	CHOWDHURY ET AL. [32]	✓	✓	✗	✗
2024	QIN ET AL. [10]	✓	✓	✓	✗

2024	TIAN ET AL. [34]	✔	✔	✕	✕
2025	LI ET AL. [35]	✔	✔	✕	✕

Distinctions in comparison: P (plot of both curves or tabular values), I (interpretation of the curves), C (correlation with wear effects)

4. Results and Discussion

Overall, the number of publications dealing with the topic of wear volume is manageable, at 36. Nevertheless, the analysis of the 25 publications identified in the structured literature search and applied to a concept matrix with four different characteristics yielded some interesting findings:

- A good half of the publications identified (56%) that deal with wear volume use the width of flank wear land as a comparative value. 16% of these use crater wear. The remaining do not use a comparative value.
- Only 48% of the publications identified compare 1D measured values with wear volume. In contrast, however, 78.6% of those that measure 1D measured values also compare it with wear volume. However, the quality of these comparisons varies greatly, which is why they are dealt with separately in Section 3.2 or Table 2.
- The analysis of the machining processes shows a clear frontrunner – turning was used in 60% of the publications. In 36% of the publications, the machining process of milling was considered, and only one publication was devoted to the machining process of drilling.
- In most publications (88% of those that specified the cutting material), carbide (68%) is used. 22.7% work with PCBN. In addition, cermet, tool steel, and HSS were each used in one publication. 76% provided information on coating. Of those who provided information, 89.5% use a coating. 26.3% worked without one. In 21% of the papers, several are examined, whereas in 10.5%, both coated and uncoated tools are compared.
- In most publications (95.2% of those that specified the material type), one material type is machined. Only one publication deals with several material types. 47.6% of those who specified a material type used a steel alloy. 33.3% of publications used titanium. In addition, one publication each used stainless steel, Inconel, CFRP, and wood.

Due to the facts that one author often published several papers about wear volume and that several authors belonged to the same research institution, the percentages shown are not valid. It goes without saying that the same measurement setups or materials were used several times.

In order to examine the research question of comparing 1D and 3D measurement more closely, Table 2 takes a closer look at this comparison. The following findings were generated from this:

- The vast majority of publications show both 1D measured values and the wear volume – but in different forms. Most show plots, while one publication presents measured values.
- Only 25% of the publications show approaches to interpretation.
- In the study by BURGER ET AL., it is noted that the wear volume curves do not align with wear curves derived from 1D parameters (e.g., VB), indicating that further research is required to clarify the relationship between 1D and wear volume [17]. The validation process of KUTTOLAMADOM ET AL. included both a comparison of traditional wear parameters with wear volume and the analysis of microphotographs. The results showed that, for turning Ti-6Al-4V, the three-dimensional method delivers wear curves comparable to those obtained using conventional approaches. [18] KUTTOLAMADOM ET AL. did not show any correlation with the wear effects listed above in their publications, but they did make a comparison with the introduced "M-ratio", which is calculated from the quotient of the volume of material machined and the wear volume [21].
- In addition to VB, ČERČE ET AL. also show the course of KT. Besides the total tool wear volume, it is also divided into crater wear volume and nose wear volume [20].

- CAPASSO ET AL. concluded that prediction based on wear volume is more accurate than prediction based on flank wear [26].
- BOING ET AL. compared the tool wear curve based on wear volume with the tool wear curve based on VB. The wear volume curve showed a constant slope and thus deviates from the characteristic curve of the VB wear curve. The authors mention surface roughness but neither elaborate nor correlate it with VB, VV, or wear effects. [27]
- YU ET AL. evaluated rake wear for different micro-texture configurations using the crater width KB. To achieve a more comprehensive characterization, wear volume and volume of adhered material were also analyzed. Although overall trends were similar, noticeable discrepancies between the graphs were not analyzed by the authors. [29]
- In the investigations by WANG ET AL. of the wear mechanisms of coated PCBN tools during high-speed turning of superalloy GH4202, wear volume served as a supplementary reference parameter but was not examined in depth. [31]
- CHOWDHURY ET AL. assessed the performance of various CrN coatings using VB, wear volume, and volume of adhered material as evaluation metrics. Although trends and rankings were similar, differing curve progressions reflected distinct wear phenomena captured by each metric. [32]
- QIN ET AL. compared wear volume with the width of flank wear land VB and observed similar overall trends, while wear volume showed higher sensitivity to variations in workpiece surface roughness. This indicates that wear volume represents tool wear progression and its effects on machining quality more accurately than classic wear parameters. [10]
- In their investigation on three-dimensional wear state detection of end mills, TIAN ET AL. compared width of flank wear land VB and wear area with wear volume. They observed that the slopes of VB and wear area curves were relatively smooth, whereas wear volume exhibited a pronounced change in slope. The rate of change of the wear volume increased significantly once rapid wear occurred. The authors concluded that wear volume can be suitable for early warning of critical tool wear. [34]
- Using their presented approach for wear detection, both classic wear parameters and wear volume were determined by LI ET AL., enabling a multidimensional evaluation of tool wear. The study primarily focused on the development of the measurement system for comprehensive multi-parameter acquisition. The average width of flank wear land and the wear area were used as additional reference metrics for evaluating the new system. [35]

5. Conclusion and Outlook

This paper first presented general information on the subject of cutting tool wear and dealt with the measurement of classic wear parameters such as width of flank wear land and crater wear. It then went on to describe another parameter, wear volume, and identified publications that deal with this topic. A structured literature review was conducted for this purpose.

Overall, it can be said that the number of these publications is manageable. For further clarification, these publications were examined in more detail and additional criteria for review were listed. These were the pure presentation of classic wear parameters in comparison to wear volume in the form of plots or tables. In addition, a small number of publications interpreted the comparison of classic wear parameters with wear volume. The formation of correlations with wear effects, such as an increase in cutting force, cutting temperature, changed chip formation, and deterioration of surface quality with increasing wear, was missing in previous publications (only few sources showed surface parameters, but did not include them in the evaluations) and requires further investigation.

Furthermore, the question of whether companies replace their tools prematurely because they are unable to adequately quantify tool wear using conventional wear parameters cannot be answered conclusively and also requires further investigation.

Appendix: List of Keywords Used in the Literature Search

Primary search terms

Primary search terms concerning wear volume: “wear volume”, “volume wear”, “volumetric wear”, “three-dimensional wear”, “3D wear”

Additional measurement-related primary search terms: “3D measurement” AND wear, “optical wear measurement”, “tool surface topography” AND wear, “tool wear measurement” AND volume, “3D data” AND wear, “point cloud” AND wear

Supplementary search terms

Process-specific supplementary search terms: cutting, cut*, milling, drilling, sawing, turning, machining, tool, “cutting edge”

Supplementary search terms addressing the comparisons between one-dimensional and three-dimensional wear metrics: VB, “flank wear”, “width of flank wear land”, “wear width”, “wear depth”

References

[1] Denkena, B., Tönshoff, H. K., 2011. Spanen: Grundlagen. 3. bearbeitete und erweiterte Auflage, Berlin: Springer Berlin. ISBN: 978-3-642-19771-0.

[2] Klocke, F. 2018. Fertigungsverfahren 1: Zerspanung mit geometrisch bestimmter Schneide. 9. Auflage, Berlin: Springer Vieweg. ISBN: 978-3-662-54206-4.

[3] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S. & Yang, L. 2021. Physics-informed machine learning. Nature Reviews Physics, 3(6), S. 422–440. DOI: 10.1038/s42254-021-00314-5.

[4] Rehorn, A. G., Jiang, J., Orban, P. E. 2005. State-of-the-art methods and results in tool condition monitoring: a review. The International Journal of Advanced Manufacturing Technology, 26(7-8), pp. 693–710. ISSN: 0268-3768. DOI: 10.1007/s00170-004-2038-2.

[5] Zhou, Y., Xue, W. 2018. Review of tool condition monitoring methods in milling processes. The International Journal of Advanced Manufacturing Technology, 96(5-8), pp. 2509–2523. ISSN: 0268-3768. DOI: 10.1007/s00170-018-1768-5.

[6] Kuttolamadom, M. A., Laine Mears, M., Kurfess, T. R., 2012. On the Volumetric Assessment of Tool Wear in Machining Inserts With Complex Geometries – Part 1: Need, Methodology, and Standardization. Journal of Manufacturing Science and Engineering, Jg. 134, Nr. 5, 2012, issn: 1087-1357. doi: 10.1115/1.4007184.

[7] Astakhov, V. P. & Davim, J. P., 2008. Tools (Geometry and Material) and Tool Wear. In: Davim, J. P. Machining – Fundamentals and Recent Advances. Springer-Verlag London Limited. ISBN: 978-1-84800-212-8.

[8] International Organization for Standardization. 3685: ISO 3685: Tool-life testing with single-point turning tools, 1993.

[9] Paucksch, E. 2008. Zerspantechnik: Prozesse, Werkzeuge, Technologien. 12. vollständig überarbeitete und erweiterte Auflage, Wiesbaden: Vieweg+Teubner Verlag / GWV Fachverlage. ISBN: 978-3-8348-0279-8.

[10] Qin, S., Wang, T., Peng, Y., Gong, T., Chen, W., Wang, H., 2024. Evaluation of Smaller Milling Cutter Health Based on Volumetric Wear Parameters. International Journal of Precision Engineering and Manufacturing (2025) 26, pp. 837–848, https://doi.org/10.1007/s12541-024-01144-y.

[11] Webster, J., Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), pp. xiii–xxiii.

[12] Cooper, H.M., 1988. Organizing knowledge syntheses: A taxonomy of literature reviews. Knowledge in Society 1 (1), pp. 104–126.

[13] Salipante, Paul F., Notz, W.I., Bigelow, J.P., 1982. A Matrix Approach to Literature Reviews. Research in organizational behavior, pp. 321–348.

[14] Moon, K. S. & Ling, Z. K., 1999. Measurement of Cutting Tool Wear with Shape-From-Focus (SFF) Method. Proceedings of: The Fourteenth Annual Meeting – The American Society for Precision Engineering, pp. 75–78.

[15] Dawson, T. G. & Kurfess, T. R., 2004. Quantification of tool wear using white light interferometry and three-dimensional computational metrology. International Journal of Machine Tools & Manufacture 45 (2005), pp. 591–596, doi:10.1016/j.ijmachtools.2004.08.022.

[16] Wang, W. H., Wong, Y. S., Hong, G. S., 2006. 3D measurement of crater wear by phase shifting method. Wear, Jg. 261, Nr. 2, pp. 164–171, issn: 00431648. doi: 10.1016/j.wear.2005.09.036.

[17] Burger, U., Kuttolamadom, M., Bryan, A., Mears, L., Kurfess, T., 2009. Volumetric FlankWear Characterization for Titanium Milling Insert Tools. In ASME 2009 International Manufacturing Science and Engineering Conference, Volume 2, ASMEDC, 2009, pp. 97–104, isbn: 978-0-7918-4362-8. doi: 10.1115/MSEC2009-84256.

[18] Kuttolamadom, M. A., Laine Mears, M., Kurfess, T. R., Burger, U., Bryan, A., 2012. On the Volumetric Assessment of Tool Wear in Machining Inserts With Complex Geometries – Part II: Experimental Investigation and Validation on Ti-6Al-4V. Journal of Manufacturing Science and Engineering, Jg. 134, Nr. 5, 2012, issn: 1087-1357. doi: 10.1115/1.4007294.

[19] Liu, J. & Xiong, G. X., 2014. Study on Volumetric Tool Wear Measurement Using Image Processing. Applied Mechanics and Materials, Jg. 670-671, pp. 1194–1199, 2014. doi: 10.4028/www.scientific.net/AMM.670-671.1194.

[20] Čerče, L., Pušavec, F., Kopač, J., 2015. 3D cutting tool-wear monitoring in the process. Journal of Mechanical Science and Technology, Jg. 29, Nr. 9, pp. 3885–3895, 2015, issn: 1738-494X. doi: 10.1007/s12206-015-0834-2.

[21] Kuttolamadom, M. A., Laine Mears, M., Kurfess, T. R., 2015. The Correlation of the Volumetric Wear Rate of Turning Tool Inserts With Carbide Grain Sizes. Journal of Manufacturing Science and Engineering, Jg. 137, Nr. 1, 2015, issn:1087-1357. doi: 10.1115/1.4028129.

[22] Kuttolamadom, M., Mehta, P., Mears, L., Kurfess, T., 2015. Correlation of the Volumetric Tool Wear Rate of Carbide Milling Inserts With the Material Removal Rate of Ti–6Al–4V. Journal of Manufacturing Science and Engineering, Jg. 137, Nr. 2, 2015, issn: 1087-1357. doi: 10.1115/1.4029649.

[23] Seid Ahmed, Y., Paiva, J., Covelli, D., Veldhuis, S., 2017. Investigation of Coated Cutting Tool Performance during Machining of Super Duplex Stainless Steels through 3D Wear Evaluations. Coatings, Jg. 7, Nr. 8, p. 127, 2017. doi:10.3390/coatings7080127.

[24] Boing, D., Schroeter, R. B., de Oliveira, A. J., 2018. Three-dimensional wear parameters and wear mechanisms in turning hardened steels with PCBN tools. Wear, Jg. 398-399, pp. 69–78, 2018, issn: 00431648. doi: 10.1016/j.wear.2017.11.017.

[25] Castro, F. L., Boing, D., Schroeter, R. B., 2018. Tool Performance Assessment based on Three-Dimensional Tool Wear Rate. Procedia CIRP, Jg. 77, pp. 638–641, 2018, issn: 22128271. doi: 10.1016/j.procir.2018.08.188.

[26] Capasso, S., Paiva, J. M., Junior, E. L., 2019. A novel method of assessing and predicting coated cutting tool wear during Inconel DA 718 turning. Wear, Jg. 432-433, p. 202 949, 2019, issn: 00431648. doi: 10.1016/j.wear.2019.202949.

[27] Boing, D., Castro, F. L., Schroeter, R. B., 2020. Prediction of PCBN tool life in hard turning process based on the three-dimensional tool wear parameter. The International Journal of Advanced Manufacturing Technology, Jg. 106, Nr. 1-2, pp. 779–790, 2020, issn: 0268-3768. doi: 10 . 1007 / s00170 - 019 -04509-0.

[28] Bai, Y., Wang, F., Fu, R., 2021. A semi-analytical model for predicting tool wear progression in drilling CFRP. Wear, Jg. 486-487, S. 204 119, 2021, issn:00431648. doi: 10.1016/j.wear.2021.204119.

[29] Yu, X., Wang, Y., Lv, D., Ye, Z., Gao, Y., 2022. Study on machining characteristics with variable distribution density micro-texture tools in turning superalloy GH4202. The International Journal of Advanced Manufacturing Technology (2022) 123, pp. 187–197, https://doi.org/10.1007/s00170-022-10117-2.

[30] Peng, Y., Qin, S., Wang, T., Hu, Y., Nie, S., 2023. Volume monitoring of the milling tool tip wear and breakage based on multi-focus image three-dimensional reconstruction. The International Journal of Advanced Manufacturing Technology (2023) 126, pp.3383–3400, https://doi.org/10.1007/s00170-023-11335-y.

[31] Wang, Z., Wang, Y., Lv, D., Yu, X., Gao, Y., 2023. Performance and wear mechanisms of coated PCBN tools in high-speed turning superalloy GH4202. The International Journal of Advanced Manufacturing Technology (2023) 127, pp. 5083–5097, https://doi.org/10.1007/s00170-023-11803-5.

[32] Chowdhury, M. S. I., Bose, B., Akter, S., Veldhuis, S. C., 2024. Effect of Deposition Parameters on Micromechanical Properties and Machining Performance of CrN Coating for Wet Finish Turning of Ti6Al4V Alloy. Materials 2024, 17, 4328. https://doi.org/10.3390/ma17174328.

[33] Mbakop Nanshie, V., Aidibe1, A., Heidari, M., Jahazi, M., 2024. Impact of coating and microstructure on wear resistance of tool steels for wood cutting: a novel approach to quantification and analysis of wear-related damages. Journal of Wood Science (2024) 70:20, https://doi.org/10.1186/s10086-024-02132-6.

[34] Tian, Y., Zhao, K., Chen, Y., Zhan, Y., Wang, T., 2024. Wear state detection of the end milling cutter based on wear volume estimation. The International Journal of Advanced Manufacturing Technology (2024) 135, pp. 1809–1824, https://doi.org/10.1007/s00170-024-14384-z.

[35] Li, L., Li, B., Wei, X., Gao, Z., Lu, Z., Guo, H., 2025. Multi-dimensional tool wear detection using Non-Subsampled Shearlet Transform and enhanced pulse coupled neural network in optical microscopy reconstruction. Optics & Laser Technology 187 (2025) 112849, https://doi.org/10.1016/j.optlastec.2025.112849.

[36] Lukáš, A., Gombár, M., Sýkora, J., Sklenička, J., Fulemová, J., Hnátík, J., 2025. Influence of Tool Clearance Angle and Cutting Conditions on Tool Life When Turning Ti-6Al-4V – Design of Experiments Approach. Journal of Manufacturing and Materials Processing, 2026, 10, 15, https://doi.org/10.3390/jmmp10010015.

[37] Degner, W., Lutze, H., Smejkal, E., 2015. Spanende Formung. Theorie, Berechnung, Richtwerte. Carl Hanser Verlag, München.

Biography

Dr.-Ing. Benjamin Thorenz (*1987) holds a PhD in Engineering Science and has been a researcher at the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth. Since 2020 he has been Senior Engineer of the department and Head of the Keylab Cutting.

M. Sc. Johannes Rauch (*1994) joined the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth in 2022 after completing his master's degree in mechanical engineering at the Technical University of Munich. He is a scientist in the field of machining with a special focus on innovative solutions for milling and drilling.

Prof. Dr.-Ing. Frank Döpper (*1970) holds the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth and is division director Manufacturing and Process Engineering at Fraunhofer IPA. For more than twenty years, he worked in technology and investment management and as a site and production manager before he was appointed professor at the University of Bayreuth.

8th Conference on Production Systems and Logistics

Development of a Digital Model for an Educational Fischertechnik Production System Using PLC-Based Control and Tecnomatix Plant Simulation

Paul Molenda, Ben Becher, Jonathan Milz, Vihasith Ganga
Hof University, Alfons-Goppel-Platz 1, 95028 Hof, Germany

Abstract

The increasing complexity of automated production systems presents challenges for both industrial and educational environments. Testing, validating and optimizing control strategies directly on physical equipment is often associated with safety risks, limited availability of systems and considerable effort in terms of time and resources. In academic contexts, these constraints additionally reduce opportunities for hands-on learning, as training equipment is frequently limited and requires close supervision. To address these limitations, this research presents the development of a digital model of an educational Fischertechnik production plant controlled by programmable logic controllers (PLCs). The real plant was systematically analyzed and all relevant mechanical, electrical and control parameters were documented. Based on this, a detailed 3D representation of the system was created and integrated into Tecnomatix Plant Simulation by Siemens, where the spatial configuration, process logic and material flow behavior were implemented to reproduce realistic operational sequences and system dynamics. The digital model enables virtual commissioning and supports experimentation and modification of control strategies in a risk-free virtual environment. It increases flexibility in teaching scenarios, reduces dependency on physical hardware and allows users to explore and understand automation processes and production system behavior more efficiently.

Keywords

Virtual Commissioning; PLC-based Automation; Tecnomatix Plant Simulation; Engineering Education; Learning Factory

1. Introduction

Digital models have become an essential component in the design, analysis, and optimization of modern automated production systems [1] . They enable early insights into system behavior, support the evaluation of layouts and process configurations, and allow control strategies to be tested before physical commissioning takes place [2]While educational production systems offer effective hands-on learning, their digital representations often lack the required functional fidelity. To address this, this study focuses on the development of a high-fidelity digital model of a Fischertechnik assembly system that reproduces the exact PLC sequencing logic to enable realistic virtual commissioning. To address this challenge, the present study focuses on the development of a detailed digital model of a Fischertechnik-based assembly system used for teaching automation and manufacturing principles. The physical installation performs a complete workflow that includes drilling, milling, assembly, packaging, and storage of a small table product. By integrating geometric data, movement parameters, and the full sequencing logic extracted from the PLC program, the

DOI: https://doi.org/10.15488/20983
ISSN: 2701-6277

digital model reproduces the behavior of the real system in a transparent and functionally consistent manner. The virtual representation supports both educational use and analytical applications, enabling learners and researchers to study operational dynamics and explore alternative process configurations through adjustable parameters.

This paper presents the conceptual foundation, modeling methodology, and implementation of the digital system, followed by a discussion of its applications in teaching and process analysis. The resulting model establishes a foundation for extending simulation-based learning scenarios and supports future research on virtual commissioning and digital representations of educational production systems.

Based on the identified limitations of existing educational models, the hypothesis of this research is that a high-fidelity digital model, which directly mirrors the step-chain logic of a real PLC rather than using simplified behavioral events, enables a significantly more accurate representation of system dynamics and station synchronization.

2. State of the art

To establish a comprehensive theoretical foundation, this section reviews the current research landscape regarding digital models, the principles of virtual commissioning, and the application of simulation tools within modern engineering education.

2.1 Digital models in production systems

Digital models play a central role in analyzing and optimizing production systems by providing virtual representations of material flows, machine behavior, and control logic. They enable engineers to evaluate process configurations, identify bottlenecks, and assess system performance before implementing physical changes. As highlighted by GRIEVES & VICKERS [6], digital representations serve as a key element of cyber-physical production systems, facilitating the exchange of information between the physical and virtual domains. Their use helps reduce development time, improve system reliability, and support structured experimentation without interrupting ongoing operations [7].

2.2 Virtual commissioning and model-based engineering

Virtual commissioning (VC) enables the testing and validation of automation logic within a simulated environment before a physical system is brought into operation [5]. By connecting virtual representations of machines, sensors, and actuators with PLC control logic, VC allows engineers to examine step sequences, timer structures, interlocking conditions, and error-handling routines early in the development process [8]. This reduces commissioning time, minimizes implementation risks, and supports systematic debugging of both mechanical and logical interactions [9].

In industrial settings, VC is widely used to evaluate the behavior of complex automation systems, verify safety logic, and ensure that material flows and machine movements operate as intended. It also enables scenario-based testing, such as evaluating alternative process strategies or examining system responses to disruptions, without affecting real production [10]. Although VC is well established in industry, its integration into educational environments remains limited. Many teaching systems rely on simplified or incomplete models that do not reflect real sequencing behavior or the complexity of PLC-controlled processes. Incorporating VC concepts into educational production systems offers significant benefits by allowing students to test control logic, analyze station interactions, and understand the relationship between digital models and physical automation. This provides a valuable platform for developing practical engineering competencies in a controlled and risk-free environment [11].

2.3 Educational production systems and simulation tools

Educational production systems and learning factories provide structured, practice-oriented environments for teaching manufacturing, automation, and systems engineering. These platforms enable students to engage directly with core elements of production technology, including material flow behavior, station interactions, and the fundamentals of PLC-based control. By offering hands-on experience within a controlled setting, such systems bridge the gap between theoretical instruction and practical application [12].

Digital models have become increasingly important within these educational environments. They complement physical installations by allowing learners to visualize internal processes that are difficult to observe directly, such as step-chain logic, sensor-actuator coordination, and the timing dependencies between stations [3]. Virtual representations also allow for experimentation with alternative process configurations or parameter settings without disrupting the physical system. This strengthens conceptual understanding and supports analytical skills through simulation-based exploration.

A critical limitation in many current educational models is the lack of detailed interlocking mechanisms and station synchronization. By focusing on these elements, digital models can bridge the gap between simplified simulations and complex industrial requirements.

Despite these advantages, many existing educational models lack the fidelity required to capture realistic control behavior or dynamic system responses [4]. Simplified digital abstractions often omit critical elements such as interlocking mechanisms, release conditions, or station synchronization, limiting their usefulness for teaching advanced automation concepts [13]. Developing accurate and functionally consistent digital models is therefore essential for enabling students to experience realistic system behavior and for supporting more sophisticated learning scenarios that involve process analysis, virtual commissioning, and parameter-driven experimentation.

3. Research gap

Unlike established virtual commissioning (VC) solutions that often rely on simplified black-box behaviors, this approach provides a 1:1 mapping of the original PLC GRAPH and FUP logic into the simulation environment. This enables students to identify specific logical errors in the control code—such as interlocking conflicts or timing-sensitive transition conditions—which are typically obscured in abstracted digital twin models.

To identify the current state of research on the use of digital models and virtual commissioning in engineering education, a systematic literature review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The databases Google Scholar, IEEE Xplore, and ScienceDirect were searched using combinations of the terms Learning Factory, Engineering Education, Virtual Commissioning, and Automation. Title-level filtering was applied to avoid the high number of irrelevant results typically produced by full-text searches. The term simulation was intentionally excluded, since it often yields broad results unrelated to virtual commissioning, while still allowing relevant studies to be captured due to the inherent simulation component of VC.

The search process initially identified 123 publications. After removing duplicates (n = 22), 101 records remained. These publications were screened based on title, abstract and full-text availability, as well as recency and relevance. A total of 100 records were excluded in this phase. Ultimately, only one publication was identified as directly related to the research focus: the study by CHAU ET AL. [14] on a digital learning factory for automation education, which demonstrates the use of virtual replicas and digital engineering tools for teaching PLC programming and system behavior.

The complete filtering process is illustrated in Figure 1, which summarizes all stages of the PRISMA workflow. The limited number of relevant studies indicates a clear research gap. While the identified publication highlights the value of digital learning environments, it does not address the development of a

high-fidelity digital model that reproduces the complete mechanical and logical behavior of a PLC-controlled educational production system. In particular, there is a lack of research on (i) the systematic extraction and transfer of real PLC sequencing logic into a simulation environment, (ii) the integration of detailed geometric and timing parameters to mirror operational dynamics, and (iii) the use of such models for virtual commissioning in an educational context. This gap underscores the need for research that develops and evaluates accurate digital representations of educational production systems to better prepare students for industrial automation practices.

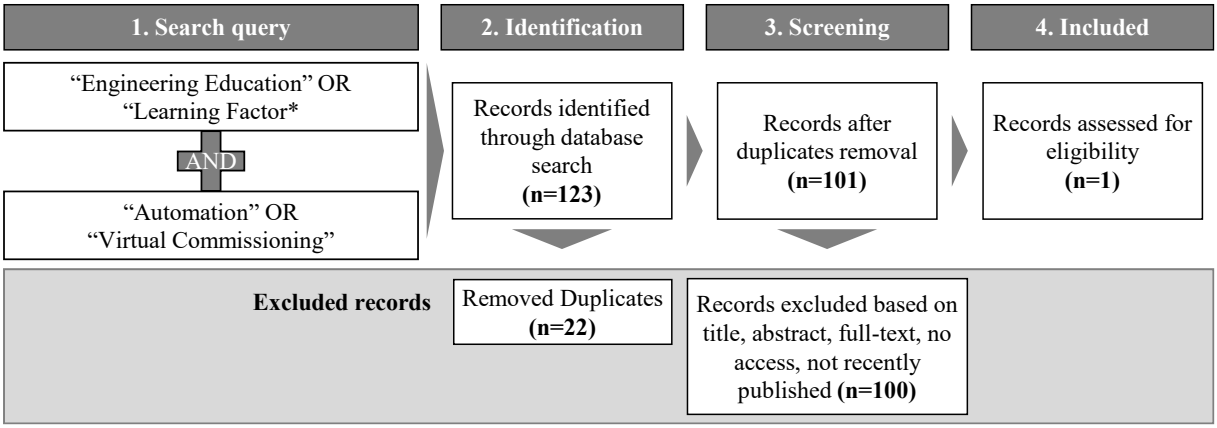


Figure 1: PRISMA flow diagram of the literature review

4. Research approach

The research approach followed a structured and reproducible workflow designed to transfer the behavior of the physical Fischertechnik assembly system into a functionally consistent digital model. The three phases of this workflow are illustrated in Figure 2 and form the methodological backbone of the modeling process.

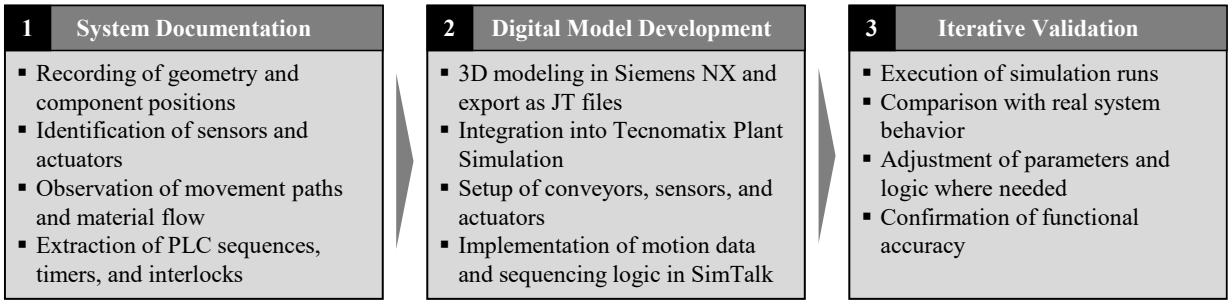


Figure 2: Overview of research approach

The research approach follows a structured three-phase workflow: system documentation, digital model development, and iterative validation (see Figure 2). The specific technical implementation and the logic mapping for the Fischertechnik system are detailed in Section 5.

5. System description

This section outlines the systematic transformation of the physical Fischertechnik assembly line into a functionally consistent digital model, encompassing the initial system analysis, the technical development in Tecnomatix Plant Simulation, and the subsequent validation of operational dynamics

5.1 System analysis and documentation

The physical Fischertechnik assembly system shown in Figure 3 serves as the reference installation for the development of the digital model. The miniature production line performs operations on a table product and provides a controlled environment in which mechanical processes, transport operations and PLC-driven sequencing can be observed in detail. To ensure that the digital representation reproduces the behavior of the real system accurately, the installation was documented comprehensively using a structured analytical procedure.

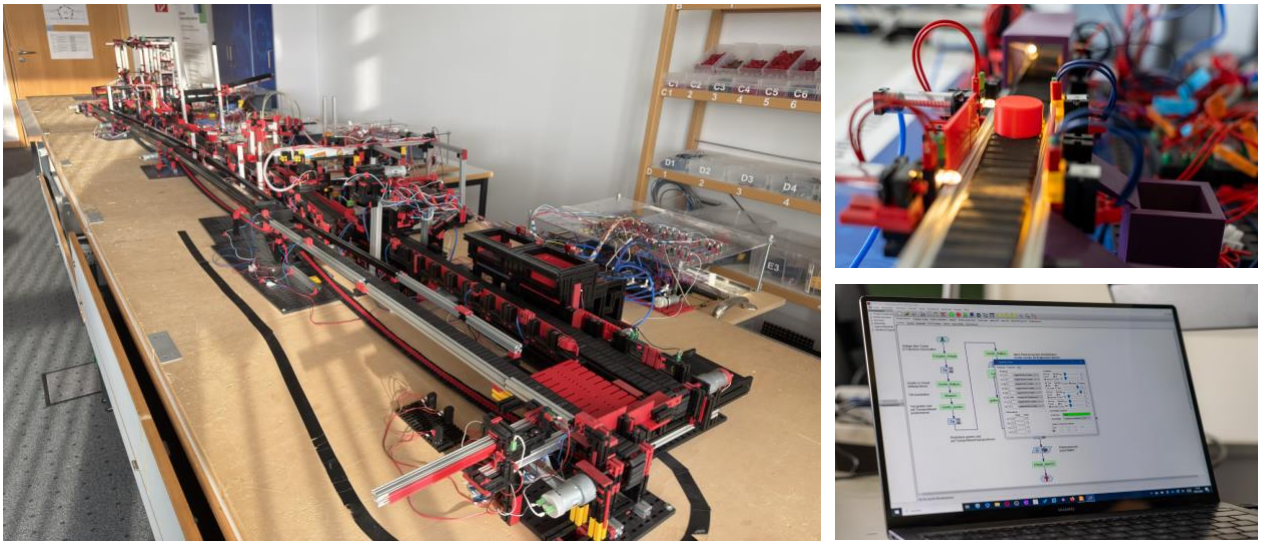


Figure 3: Physical Fischertechnik production system representing a fully automated table assembly line.

The geometric characteristics of the system were captured through detailed measurements of conveyor lengths, component dimensions, stroke distances of movable elements and the spatial arrangement of fixtures and tools. The circulation of the loading carrier was identified as the critical element for material flow stability; therefore, its transfer points and station interaction zones were prioritized during analysis and subsequent validation. The resulting geometric dataset forms the basis for the subsequent 3D modeling and layout integration.

The control logic of the physical system (PLC program), programmed using Function Block Diagram (FBD/FUP) for interlocking conditions and Sequential Function Chart (SFC/GRAPH) for step-chain sequences, was systematically analyzed to ensure a functionally consistent transfer into the simulation environment. All sensors and actuators were identified and mapped to their operational roles, including their dependencies, activation conditions and interactions with the transport system. Release and interlocking conditions, timer structures and step transitions were documented to ensure that the logical structure governing the real system could be fully reproduced in the simulation environment. This analysis provides the functional framework that defines how stations coordinate their activities and how material flow is controlled. To complement the mechanical and logical documentation, controlled test runs were conducted. During these runs, the execution of all five processing stations was observed, with attention paid to timing relationships, synchronization between stations and the interactions between processing steps and the loading carrier. These observations allowed the identification of characteristic cycle times, switching delays and behavioral patterns that cannot be derived from static documentation alone. They provide essential empirical data for calibrating the timing and movement parameters of the digital model.

The consolidated dataset resulting from geometric measurements, PLC logic extraction and process observations provides a comprehensive representation of the real system. Table 1 summarizes the main processing steps and their associated PLC logic elements, offering an overview of the system behavior

documented in this phase and forming the foundation for the subsequent modeling activities described in the following sections.

Table 1: Compact overview of the Fischertechnik table assembly line

No.	Station	Process step	Key PLC logic functions
1	Drilling Unit	Drilling holes into the table plate on the loading carrier	Drill cycle start, drill position check, safe drilling sequence
2	Milling Unit	Simulated milling of contours and edges	Milling sequence control, tool movement positioning, cycle timing
3	Assembly Unit	Assembly of legs, cross beams, and longitudinal beams	Structured assembly sequence, component presence check, assembly verification
4	Packaging Unit	Packaging of the finished table	Packaging trigger, item completion check, carrier–table separation
5	Storage Unit	Automated storage of finished tables	Slot allocation, lift positioning, buffer control

5.2 Model Development in Tecnomatix Plant Simulation

The digital model of the Fischertechnik assembly system (cf. Figure 4) was developed in Siemens Tecnomatix Plant Simulation based on the mechanical and logical characteristics documented in Section 5.1. The objective of this phase was to recreate both the physical structure and the operational behavior of the system in a consistent and transparent manner.

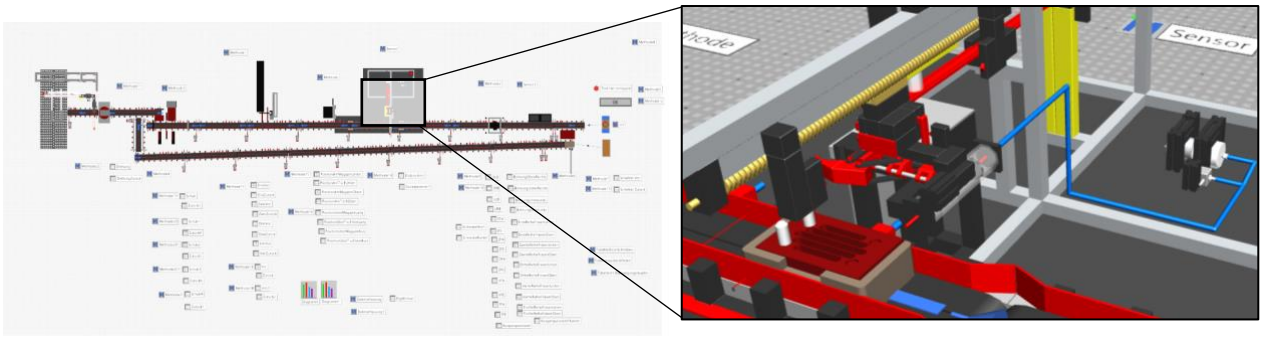


Figure 4: Digital simulation model of the Fischertechnik table assembly line in Tecnomatix Plant Simulation.

All relevant mechanical components were created in Siemens NX. Movable parts such as pushers, stoppers, the suction gantry and the drilling and milling units were modeled as separate CAD (Computer-Aided Design) bodies to enable independent animation within the simulation environment. After the CAD work was completed, the assemblies were exported as JT files and imported into Tecnomatix Plant Simulation. Each imported component was assigned to its corresponding station to establish a structured representation of the physical system. Within Plant Simulation, conveyor paths, sensors and actuators were defined and linked to the imported geometry. This parameterization supports flexible experimentation and maintains transparency of system behavior.

The model includes all stations, the circulating loading carrier and the kinematic behavior of the movable components. The resulting visual representation provides the environment for implementing and testing control logic. The sequencing logic of the real system was implemented using SimTalk. The step chains defined in GRAPH and the logical conditions implemented in FUP were transferred into SimTalk routines that govern the order of process steps, coordinate handovers between stations, manage timer structures and ensure synchronized movements. By mapping the PLC logic into the simulation environment, the digital model consistently reflects the functional behavior of the physical system. The data sources used for model

development, including geometric measurements, PLC logic and NX-based 3D models, are summarized in Table 2. These elements form the foundation for achieving an accurate digital representation of the Fischertechnik assembly system.

Table 2: Overview of data sources and control logic components

Data source or logic component	Content and purpose
Manual geometric measurements	Dimensions of stations, conveyor lengths, sensor locations and movement distances
Manual process-time measurements	Cycle times, transport durations, suction times, switching delays
PLC program in FUP and GRAPH	Step sequences, timer values, release conditions and interlocks
NX 3D models exported as JT-file format	Representation of all static and movable components for animation
Motion databases	Storage of movement parameters including positions, angles, velocities and timing values
SimTalk process methods	Control of station operations and coordination of complete sequences
External spreadsheet tables	Parameter input for timing and movement values without code modification

5.3 Implementation and validation of process behavior

The validated behavior of the Fischertechnik assembly system was examined directly within the simulation model shown in Figure 4. Based on the sequencing logic and motion definitions, the objective of this phase was to verify that the simulated processes match the operational characteristics of the physical installation. Validation was carried out through a series of structured simulation runs. These runs focused on assessing material flow continuity, the correct timing and ordering the stations, as well as the synchronization between stations.

The simulated behavior was compared with observations recorded during the analysis of the physical system. Discrepancies, such as timing deviations, incorrect sensor responses or misaligned movement sequences, were systematically corrected by adjusting parameter tables, refining sensor placements or modifying SimTalk routines. This iterative refinement ensured that the digital model reproduced the relevant functional dependencies of the real system. Through this validation process, the model achieved a consistent and reliable representation of system behavior. It provides a robust foundation for the results and application scenarios discussed in Chapter 6 and supports both instructional and analytical use.

6. Results and application

The validated simulation model provides an accurate and consistent representation of the Fischertechnik assembly system. All major process steps follow the sequencing logic extracted from the PLC program, and the parameterized motion data ensure realistic timing and movement behavior. Material flow, carrier circulation, and station interactions are reproduced reliably. A key result is the ability to modify process parameters without changing SimTalk routines. Adjusting speeds, timing values, or release conditions through external tables allows users to analyze the effects of configuration changes and to study system performance under varying conditions. The model supports university teaching by visualizing PLC controlled processes, step chains, and sensor actuator interactions. The 3D animation helps students understand mechanical behavior that is not easily observable on the physical system. Its applicability across several teaching formats is summarized in Table 3.

Table 3: Use of the digital model in university education

Educational area	Application	Learning outcomes
Automation Engineering	Virtual testing of PLC logic and sequences	Understanding of control structures and troubleshooting
Simulation and Manufacturing Systems	Parameter studies and material flow analysis	Skills in simulation modeling and system evaluation
Virtual Commissioning	Soft PLC integration for virtual tests	Competence in VC workflows and error analysis
Industry 4.0	Demonstration of digital system behavior	Understanding of cyber physical systems
Lean Production	Analysis of cycle times and bottlenecks	Process optimization capabilities
Mechatronics	Visualization of integrated system behavior	System level understanding
Student Projects	Model extensions and scenario experiments	Applied problem solving

Beyond education, the model enables scenario analysis and virtual commissioning tasks, including testing alternative control strategies and evaluating process changes without affecting physical equipment. These capabilities demonstrate its suitability for instructional and analytical use and provide a basis for future extensions.

7. Summary and outlook

This paper presented the development of a detailed digital model of a Fischertechnik assembly system using Tecnomatix Plant Simulation. This paper demonstrated that high-fidelity digital modeling based on direct PLC logic extraction significantly enhances the functional accuracy of educational simulations. The resulting system serves as a robust tool for both advanced automation training and virtual commissioning research.

In addition, the model supports simulation-based experimentation by enabling users to adjust process parameters and evaluate the resulting system behavior.

Future development efforts will focus on extending the digital model to include additional components of the real installation. This includes integrating the high-bay storage system with its crane movements and slot allocation logic, as well as implementing the nonconformance handling process that directs rejected products back to the beginning of the line. Enhancing the parameter table system and adding advanced analysis features will further improve usability and support more sophisticated simulation studies. Another promising direction is the coupling of the digital model with a soft PLC to enable comprehensive virtual commissioning scenarios and to provide an expanded environment for testing and education.

References

[1] Banks, J., Carson, J., Nelson, B.L., MacNicol, D., 2014. Discrete-event system simulation, Fifth edition ed. Pearson, Harlow, United Kingdom, 1560 pp.

[2] Law, A.M., 2015. Simulation Modeling and Analysis, Fifth Edition ed. McGraw-Hill US Higher Ed USE Legacy, New York, 776 pp.

[3] Zuehlke, D., 2010. Smart Factory - Towards a factory-of-things. Annual Reviews in Control 34 (1), 129–138.

[4] Harrison, R., Vera, D., Ahmad, B., 2016. Engineering Methods and Tools for Cyber–Physical Automation Systems. Proc. IEEE 104 (5), 973–985.

[5] Oppelt, M., Urbas, L., 2014. Integrated virtual commissioning an essential activity in the automation engineering process: From virtual commissioning to simulation supported engineering, in: IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society. Dallas, TX, USA, 29 October – 1 November 2014. IEEE, Piscataway, NJ, pp. 2564–2570.

[6] Grieves, M., Vickers, J., Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems.

[7] Kagermann, H., Wahlster, W., Helbig, J., 2013. Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0.

[8] Siemens AG, 2019. SIMIT Simulation Platform.

[9] Lee, J., Bagheri, B., Kao, H.-A., 2015. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters 3, 18–23.

[10] Hoffmann, P., Schumann, R., Maksoud, T.M.A., Premier, G.C., Virtual commissioning of manufacturing systems.

[11] Lechler, T., Fischer, E., Metzner, M., Mayr, A., Franke, J., 2019. Virtual Commissioning - Scientific review and exploratory use cases in advanced production systems. Procedia CIRP 81, 1125–1130.

[12] Abele, E., 2019. Learning factories: Concepts, guidelines, best-practice examples. Springer, Cham.

[13] Bolton, W., 2015. Programmable logic controllers, Sixth edition ed. Newnes, Kidlington, Oxford, UK, Waltham, MA, 11 pp.

[14] Chau, C., Su, V., Tran, K., Nguyen, C., 2025. Digital Engineering Approach for Automation Education: A Case Study in Using Digital Learning Factory. In: Decarbonizing Value Chains. Springer, Cham, pp. 772–781.

Biography

Paul Molenda (*1988) is a Professor of Industrial Engineering at Hof University of Applied Sciences, where he focuses on the combination of digitalization, lean production, and resource efficiency. He leads the Smart Factory Lab at Hof University of Applied Sciences and is actively involved in research and teaching. He earned his doctorate at the University of Bayreuth and previously worked as a project manager at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA and as the Head of Digital Transformation at ZF Friedrichshafen AG.

Ben Becher (*2002) is a Bachelor student of Industrial Engineering at Hof University of Applied Sciences, specializing in Digital Factory concepts with a focus on AI and digitalization. He is part of a cooperative study program with REHAU Industries SE & Co. KG. His academic interests include the design and optimization of production systems, as well as the analysis, modeling and simulation of manufacturing processes using digital models.

Jonathan Milz (*2002) has been studying Engineering Sciences at Hof University of Applied Sciences since 2022 and is currently in his seventh semester. His program focuses on Digital Factory and covers the planning, design and optimization of industrial production systems. His main interests include digital-twin concepts as well as the simulation and analysis of material flows and assembly processes.

Vihasith Ganga (*1998) is a master's student in Engineering and Business Administration at Hof University of Applied Sciences, specializing in operational excellence. He is a student assistant at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Bayreuth, and his research areas include production systems, digitalization, and remanufacturing.

8th Conference on Production Systems and Logistics

Development And Evaluation Of An AI-Enhanced Smart Glove For Scannerless Order Picking In Intralogistics

Jan Schuhmacher[1], Vera Hummel[1,2]

[1]*Reutlingen University, NXT Sustainability and Technology, Alteburgstraße 150, 72762 Reutlingen, Germany*
[2]*Department of Industrial Engineering, Stellenbosch University, 7600 Stellenbosch, South Africa*

Abstract

This paper presents the development and evaluation of an AI-enhanced smart glove system designed to enable scannerless execution of pick and stow operations in industrial warehouse environments. The core innovation lies in the combination of a wearable device that captures precise hand movements and spatial positions via an indoor localization system, and a novel AI-based method for clustering and combining pick and stow orders. This integration allows an automated detection and confirmation of material handling operations based on walking distance optimized pick and stow order packages without manual barcode scanning or user input. Real-time motion and sensor data from the glove is processed using machine learning algorithms to identify pick and stow activities and detect errors based on positional discrepancies. Detected anomalies are immediately communicated to the operator through acoustic feedback. In parallel, incoming pick and stow orders from ERP systems are dynamically clustered and transformed into optimized task packages. These are autonomously allocated to the nearest available operator, based on spatial proximity and system workload, enabling self-directed and efficient task execution. The system is tested and validated within Werk150, the research and learning factory at Reutlingen University, which replicates realistic assembly and intralogistics processes for variant-rich low-volume production. The work has been conducted as part of the collaborative research project "ALISA – AI- and Localization-Based Order Picking System with Intelligent, Scannerless Work Glove" involving leading small and medium-sized enterprises (SMEs) and academic institutions specializing in logistics and resilient systems. The project focused on integrating intelligent wearables with AI-based decision-making to increase responsiveness and adaptability in intralogistics. The results achieved confirm the feasibility and effectiveness of the proposed approach. Replacing conventional scanning systems with intelligent, AI-supported wearables leads to measurable improvements in process efficiency, error reduction, and operational flexibility in dynamic warehouse environments.

Keywords

Intralogistics; Artificial Intelligence; Smart Wearable; Indoor Localization; Process Optimization

1. Introduction

The production and market environment of manufacturing companies is characterized by high volatility, driven by personalized products and small batch sizes [1]. In particular, small and medium-sized enterprises (SMEs) often operate in niche markets with low production volumes, while facing significant cost pressures and stringent product quality requirements as key differentiators [2]. Across various industries, logistics costs account for approximately 25% to 45% of total product manufacturing costs, with a substantial portion attributed to warehousing and order picking [3]. Due to the diversity of handled products and the required

DOI: https://doi.org/10.15488/20984
ISSN: 2701-6277

flexibility, these processes are predominantly manual. Picking is generally the most labor-intensive activity within warehouse operations, representing up to 60% of all tasks [4,5]. Consequently, its optimization is critical for improving logistical performance in terms of goods processed per unit time.

However, traditional centralized, sequential planning and optimization approaches based on historical and experiential data are no longer effective in reducing picking times (e.g., through zone-based strategies), given increasingly shorter change cycles in product lifecycles and demand patterns [6]. Manual workflows, including barcode scanning and other feedback steps, remain time-consuming, error-prone, and inflexible, while RFID-only solutions require dedicated stations and fixed transport paths, limiting adaptability [4].

Significant potential exists for cost savings and process optimization through resilient location-based technologies combined with AI-driven analytics, enabling real-time monitoring and adaptive optimization within SME ERP systems.

To address this research gap, the ALISA project served as the design and evaluation context for developing and validating a novel scannerless order picking approach. The collaborative research project involved industrial partners (providing work system infrastructure and industrial gloves) and academic institutions. The wearable smart glove captures hand movements and spatial positions via indoor localization, while machine learning algorithms identify pick and stow operations and detect errors, providing immediate feedback. ERP orders are dynamically clustered into walking-distance-optimized packages and allocated to the nearest operator, supporting efficient self-directed execution. The ALISA system therefore constitutes the prototype artefact developed and evaluated following a design science research (DSR) methodology. The system was implemented and validated at Werk150, Reutlingen University's research and learning factory.

2. Research method

The research adopts a DSR methodology [7] with a problem-centered initiation of the research process that combines iterative system development with empirical validation in a realistic intralogistics environment. Within this methodological framework, the ALISA system represents the artefact designed to address the identified inefficiencies in manual order picking processes. Research activities were carried out within the ALISA project, involving industrial partners who provided the operational requirements and wearable hardware, and academic partners who contributed methodological and analytical expertise.

The ALISA system integrates three technical components into a unified scannerless picking workflow: a wearable smart glove with knitted-in resistive fingertip sensors for detecting gripping actions and hand movements; an ultrasonic indoor localization system that captures high-resolution spatial trajectories of the operator's hand; and an AI-based control platform responsible for real-time gesture classification, task optimization, and operator feedback. Together, these components enable the automated detection and confirmation of pick and stow operations without manual barcode scanning.

Data acquisition involved continuously capturing fingertip resistance values, motion data, and spatial coordinates from the localization system. These data streams were synchronized with live order information retrieved from the ERP system. A machine learning pipeline processed the combined sensor data to classify operator activities into pick, stow, or idle states, using features extracted from resistance patterns, movement trajectories, and spatial context. Error detection was implemented through a rule-based comparison of expected and observed glove sensor and location data. Detected deviations triggered immediate acoustic alerts to the operator.

To enhance operational efficiency, incoming pick and stow orders were dynamically clustered into walking-distance-optimized task packages using AI-based methods. The allocation algorithm considered the spatial proximity of items, real-time operator locations, and current workload distribution. Task packages were then autonomously assigned to the most suitable operator, enabling balanced, self-directed execution and eliminating the need for manual confirmation steps.

Experimental validation of the system was conducted in Werk150, Reutlingen University's research and learning factory. This environment replicates realistic intralogistics processes for variant-rich, low-volume production and provides controlled yet representative conditions for evaluating system performance. Validation focused on the accuracy of activity classification, the reliability of error detection, operator responsiveness to feedback, and the efficiency gains achieved through dynamic task clustering and autonomous task allocation.

3. Analysis of the State-of-the-Art

This chapter provides a structured analysis of the current state-of-the-art in science and technology relevant to warehouse operations. The first part examines research on AI-based clustering of pick-and-stow orders, highlighting approaches that aim to optimize efficiency and accuracy in order handling. The second part focuses on the provision of information and the capturing of process data, with particular attention to technologies such as smart gloves that enable monitoring and feedback during operational tasks. Together, these perspectives establish the conceptual and methodological foundation for the research contributions presented in the subsequent sections of this paper.

3.1 AI-based clustering of pick and stow orders

A comprehensive review of the current scientific and technological state-of-the-art was conducted, focusing on AI-based methods for optimizing intralogistics processes, particularly clustering of pick and stow operations. The analysis included an evaluation of existing algorithms and approaches to identify the most suitable techniques for integration into the ALISA system.

Three major categories of algorithms were examined and the detailed analysis has been published in Schroth et al. [8]:

- **Evolutionary Algorithms and Metaheuristics (EAM)**: Methods such as Genetic Algorithms and Particle Swarm Optimization are widely applied for optimizing picking operations in 2D warehouse layouts and, to a lesser extent, stowing processes. These approaches typically rely on extensive historical datasets [9–12].
- **Heuristic Algorithms (HA)**: Techniques including heuristic-based methods and Simulated Annealing aim to accelerate solution finding through shortcuts and domain-specific insights. They are primarily used for clustering pick operations in 2D layouts [13–16].
- **Clustering Algorithms (CA)**: Algorithms such as k-means and its variants are employed to detect patterns and categorize data points. These methods are applicable to both 2D and 3D warehouse layouts and are increasingly used for optimizing combined pick and stow operations [17–20].

The review revealed two key limitations in current research: (i) most AI approaches rely heavily on historical data, neglecting real-time information, and (ii) pick and stow processes are often optimized in isolation, complicating their integration into holistic intralogistics systems.

To address these gaps, several clustering algorithms were evaluated within the context of the ALISA project using a practical scenario involving four pick positions. Among the evaluated methods, only k-means and Agglomerative Clustering correctly identified the optimal number of clusters. Based on performance and suitability for dynamic, real-time environments, k-means was selected as the clustering algorithm for the ALISA system.

For autonomous allocation of bundled pick-and-stow tasks, an agent-based approach was implemented, building on prior research and extended methodologies [21]. Further analysis during the project confirmed that clustering algorithms, particularly k-means, offer significant potential for real-time optimization in both 2D and 3D layouts, making them a promising solution for integrated intralogistics systems.

3.2 Provision of information and capturing of process data

Information and process feedback in warehouse environments are traditionally provided through mobile handheld devices and data collection tools [22]. These systems transmit pick and stow instructions visually via displays or audibly through headsets, with task completion confirmed either by barcode scanning or verbal voice commands. Additionally, pick-by-light systems permanently installed on shelves are employed to guide operators. All these hardware solutions are typically connected to Enterprise Resource Planning (ERP) or Warehouse Management Systems (WMS), which serve as data sources and sinks [23,22]. However, these systems lack the capability to record near-real-time process or positional data, such as time or location, and rely on manual feedback from the workforce.

To address localization requirements, specific indoor positioning systems based on technologies such as Wi-Fi, Bluetooth, ultra-wideband, and ultrasound have been introduced, enabling position tracking within meter to centimeter accuracy ranges [24,25]. Despite these advancements, the integration of real-time process data collection into picking and stowing workflows remains limited.

The development of smart gloves represents a critical step toward closing this gap, as current solutions do not provide comprehensive near-real-time process data. Existing products, such as ProGlove [26], primarily focus on barcode scanning for process feedback, thereby limiting broader data capture capabilities in warehouse operations. Other glove-based solutions primarily target ergonomic work assessment by integrating sensors, wireless data transfer, and software for data export [8]. These systems enable the creation of force catalogs that support ergonomic analysis using methodologies like EAWS, based on action force and its frequency over time [27].

3.3 research gap

Despite significant progress in AI-based clustering methods and the introduction of wearable technologies for intralogistics, current research reveals two major shortcomings. First, most optimization approaches rely predominantly on historical datasets, thereby neglecting the integration of real-time process and positional information into clustering and task allocation. Second, existing wearable solutions, such as smart gloves, are primarily limited to barcode scanning or ergonomic assessments and do not enable comprehensive, near-real-time data capture of material handling activities. Moreover, pick and stow processes are often treated in isolation, which hinders their seamless integration into holistic intralogistics systems. Addressing these gaps requires the development of intelligent, scannerless wearable solutions that combine real-time motion and localization data with AI-driven clustering to enable autonomous, efficient, and error-resilient execution of pick and stow operations.

4. AI-Enhanced Smart Glove System ALISA

The ALISA system integrates wearable sensing, AI-based data processing, and real-time feedback to enable scannerless, adaptive intralogistics. The smart glove captures motion and handling data from pick and stow operations and synchronizes it with ERP/WMS information, while a standardized interface ensures seamless communication with AI modules and assistive devices. This architecture enables real-time activity detection, anomaly reporting, and automatic feedback to ERP systems without manual scanning. The following sections detail the system components, highlighting how sensors, algorithms, and interfaces work together to enhance efficiency, flexibility, and transparency in warehouse operations.

4.1 System overview

Building on the identified requirements for real-time data integration and AI-based clustering of pick and stow operations, the ALISA project advances toward a holistic system concept. At its core, the smart glove serves as the central data acquisition device, seamlessly linking manual activities with digital process control. Through integrated sensor technology, the glove captures motion and handling data, which are synchronized with master and transaction data from ERP/WMS systems. A standardized CSV-based interface ensures

robust and efficient data exchange between subsystems, enabling both the clustering of orders and the provision of task-specific information to operators via assistive devices such as smart glasses, tablets, or panel PCs (also see Figure 1).

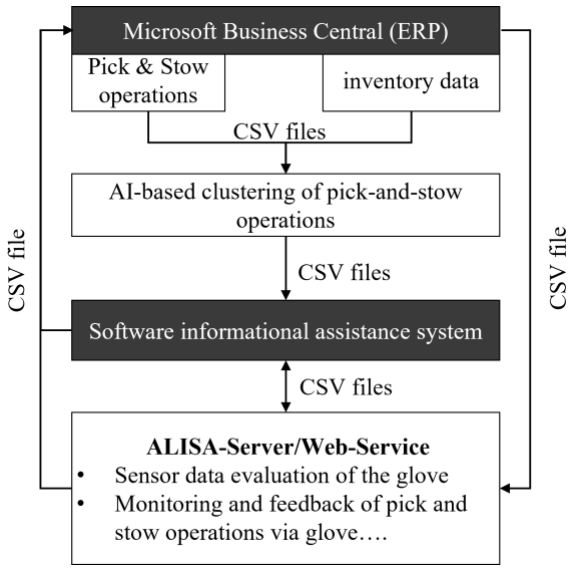


Figure 1: Interface concept

This architecture establishes the foundation for the AI-enhanced smart glove system, in which detected pick and stow activities are processed in real time, anomalies are communicated immediately, and task completion is fed back into ERP systems without manual scanning. The following section outlines the major system components, illustrating the interplay of hardware, software, and interface components that enable scannerless, adaptive, and efficient warehouse operations.

4.2 AI-based clustering of pick and stow operations

The developed solution component for the AI-based clustering of pick and stow operations addresses intralogistics optimization through a twofold approach. First, a detailed simulation model of Werk150 was implemented using AnyLogic to create a synthetic dataset for training and testing. The warehouse and adjacent modules were represented as a digital shadow in a 3D environment, incorporating Methods-Time Measurement (MTM) standards to accurately replicate manual task durations [28]. This synthetic data mitigates the scarcity of real-world datasets and enables controlled experimentation and scalability across diverse logistics scenarios.

Second, the generated data was employed to train an unsupervised machine learning model developed in TensorFlow. By applying the k-means clustering algorithm, the model identifies operational patterns and forms clusters that minimize travel distances and handling times, thereby improving overall logistics performance. A detailed description of the implementation process is provided by Schroth [8]. The optimized cluster of pick and stow operations, generated through the AI-based approach, is presented to the operator prior to execution via a Power BI dashboard (also see Fig. 3). This interface displays the planned walking distance alongside a visual representation of the corresponding pick and stow positions within the warehouse, ensuring transparency and supporting efficient task execution.

4.3 Smart glove

The ALISA smart glove is a low-cost textile work glove with embedded resistance-based sensors at the fingers and palm, connected to a wristband controller that processes the glove sensor data and captures location information of pick and stow positions using an ultrasonic-based 3D indoor localization system. Characteristic resistance patterns indicate hand movements and grasping actions (See Fig. 2). Combined with the positional data, the glove reliably detects pick and stow operations in warehouse environments.

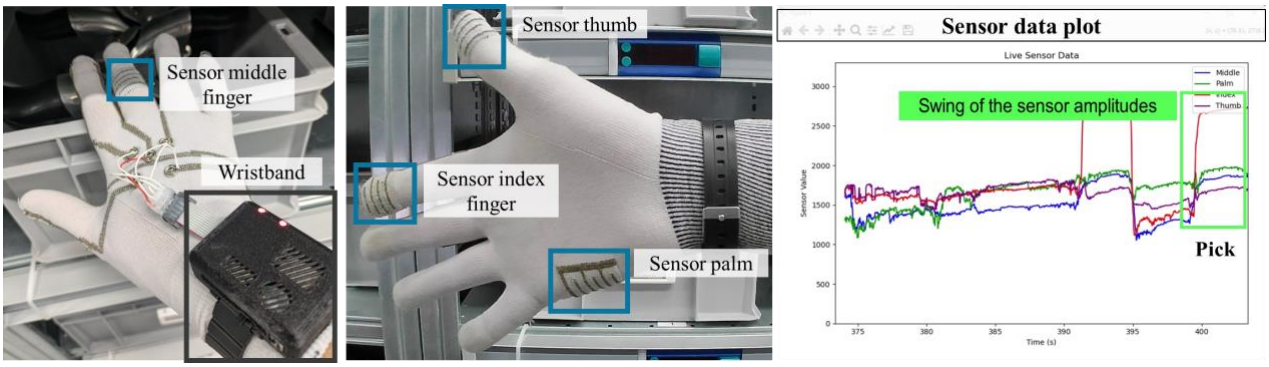


Figure 2: Smart glove and sensor amplitudes of a pick

The project expanded this capability with artificial intelligence for gesture and object class prediction. Both passivated and unpassivated glove variants were examined. Sensor fusion integrated gesture data, object labels and glove location to increase accuracy. Deep learning models including Transformer, CNN LSTM and Inception variants were trained on several object classes such as cylinder, mud guard, foot stand, mechanical fixture and tire. The Inception model reached the highest validation accuracy at 87%. By combining motion, resistance and spatial information, the system enables scannerless identification and confirmation of handling activities and provides real time monitoring, error detection and feedback to improve warehouse efficiency.

5. Evaluation of the developed solution

For the practical development, testing, and validation of the ALISA overall system with the smart glove, test and validation scenarios were designed and implemented at the warehouse area of Werk150, covering various picking and stowing configurations.

Initially, additional sensor training data for gesture recognition and object classification were generated for the research group at the University of Freiburg [29]. Subsequently, the basic functionalities of the smart glove, including gesture recognition (also see Fig. 2) and localization within the Werk150 warehouse, were tested in collaboration with project partners from the University of Freiburg [29] and the industry partner Telocate [30].

After completing these initial tests, realistic validation scenarios were developed and executed together with the industrial partner BeeWaTec, simulating different pick and stow configurations. An exemplary scenario is described below to illustrate the validation of the ALISA overall system at Werk150.

According to the developed ALISA system and interface concept (Fig. 1), the process cycle begins with the export of pick and stow operations and master data (including inventory, warehouse definition with bin designations, and X/Y/Z coordinates) from the ERP/WMS system. These data serve as input for k-means-based clustering to generate combined order packages. The resulting order packages, to be executed autonomously by the nearest warehouse staff, are displayed to logistical decision-makers through a custom Microsoft Power BI dashboard (Fig. 3). The dashboard includes a table showing the route-optimized sequence of pick and stow operations for the identified order cluster (Fig. 3, top left), a visualization of pick (red triangle icon) and stow (green square icon) positions in the racks (Fig. 3, bottom), the number and proportion of operations within the cluster (Fig. 3, top right), and the planned travel distance of the warehouse staff in meters for task completion (Fig. 3, top right). This information is updated during process execution with the captured pick and stow positions of the glove (Fig. 3, bottom, blue circle icons) and the actual walking distance.

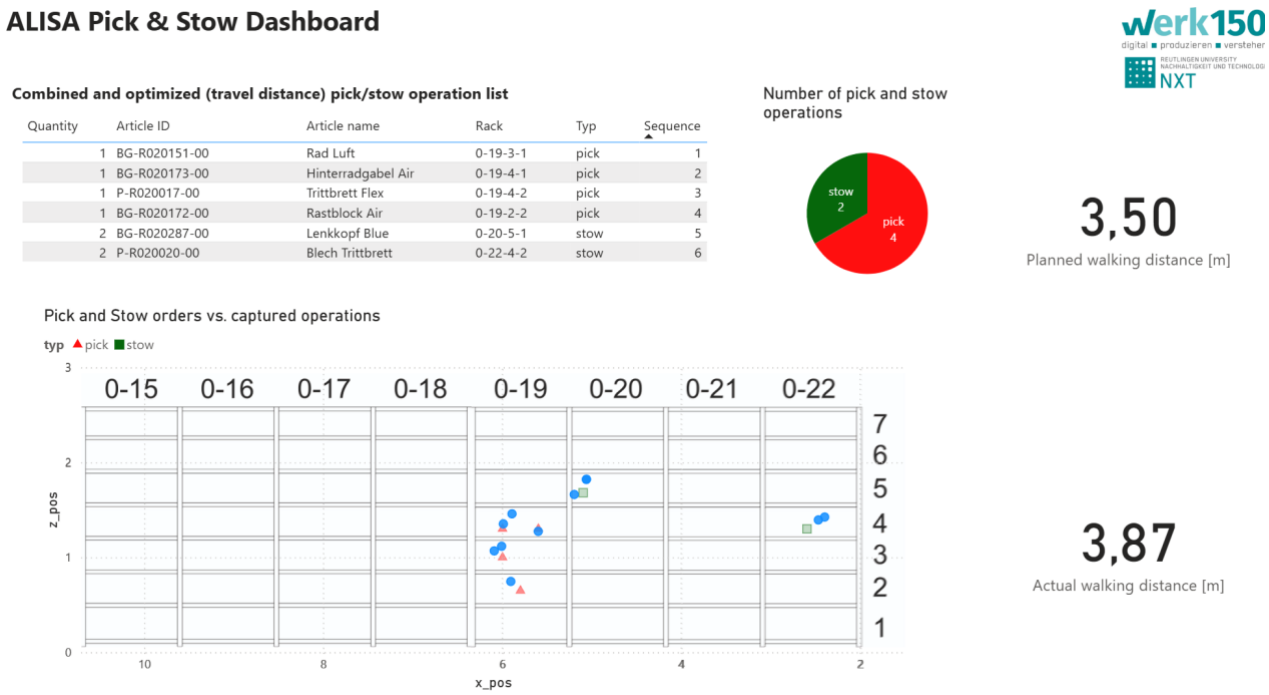


Figure 3: Microsoft Power BI Dashboard of pick and stow operations

During the execution of a combined pick and stow order, the warehouse employee receives process-relevant information, such as the current pick or stow task, storage location, item, and quantity, directly in their field of view via an informational assistance system based on Microsoft HoloLens Smart Glasses and Microsoft Guides software (Fig. 4, top left). Process feedback for completed pick and stow operations, derived from smart glove sensor data (finger sensors and localization information), can be monitored live via the ALISA portal (Fig. 4, top right & Fig. 5).

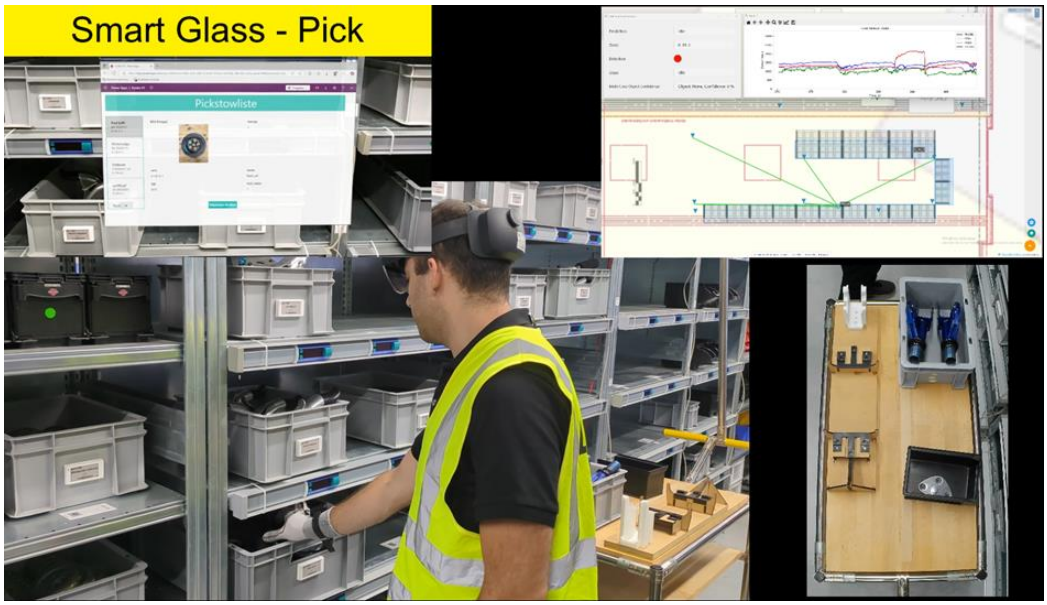


Figure 4: Validation scenario Werk150 with informational assistance (top left) and ALISA portal (top right)

The ALISA portal (Fig. 5) provides an integrated overview of the data captured by the smart glove, including amplitude signals from the finger and hand sensors (Fig. 5, top right), the glove's spatial position within the warehouse (Fig. 5, bottom), and derived information such as detected pick or stow operations, warehouse zone, and the classified object type together with its associated confidence value (Fig. 5, top left). This

information is transmitted to the overarching ERP/WMS system for process feedback, enabling real-time monitoring, error detection, and data-driven support for warehouse operations. Potential errors are detected via a rule-based comparison of expected and observed glove sensor and location data and deviations triggered immediate acoustic alerts to the operator. The validation scenarios demonstrated an activity classification accuracy of about 90% for pick, stow, and idle actions, object type classification accuracy between 60% and 80%, and a reduction in walking distance of 10% to 40% for combined pick and stow orders compared to processing them separately, alongside reliable error detection and prompt operator responsiveness to acoustic alerts.

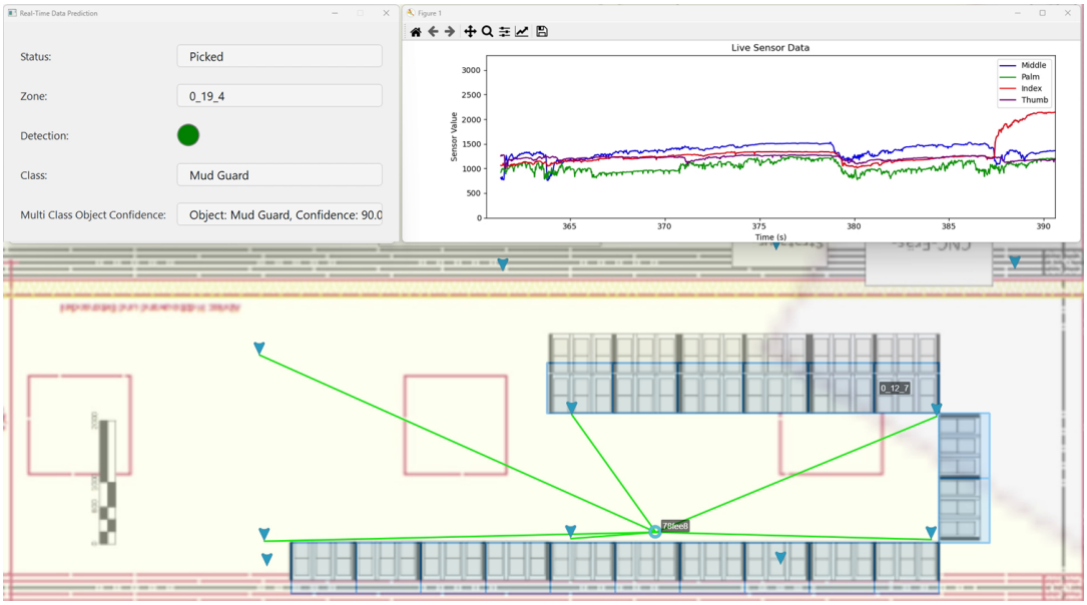


Figure 5: ALISA portal

6. Conclusion and outlook

The developed ALISA smart glove system successfully demonstrated the integration of sensor-based gesture recognition, object classification, and indoor localization for scannerless warehouse operations. Practical validation at Werk150 showed that the glove could reliably capture pick and stow operations in real time, providing both process feedback and spatial information to the ALISA portal and ERP/WMS systems. The combination of finger and palm sensors with a wrist-mounted controller enabled accurate gesture detection, while the use of deep learning models allowed for object classification, achieving high validation accuracy, particularly for the Inception model. However, object classification still requires further optimization, and additional joint sensors could enhance performance by providing more detailed curvature information. The system introduces only moderate additional hardware costs and leverages SME-standard ERP platforms such as Microsoft Dynamics 365 Business Central, supporting practical applicability for small and medium-sized enterprises. Future work will focus on integrating a near real-time interface to Microsoft Dynamics 365 Business Central for direct access to ERP/WMS data, and optimizing hardware and software for scalable, cost-effective, and market-ready use. Sensor selection will be guided by operational needs, balancing gesture recognition with object classification to complement localization data, further enhancing efficiency and usability of AI-driven intralogistics solutions.

Acknowledgements

This research has been supported by the Ministry of Economic Affairs, Labor and Tourism Baden-Wuerttemberg (Germany), as part of the "Invest BW" funding program, within the framework of the research project “AI- und Lokalisationsbasiertes Kommissioniersystem mit intelligentem Arbeitshandschuh (ALISA) (BW1_3112/05)”.

References

[1] Bauernhansl, T., 2017. Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma, in: Vogel-Heuser, B., Bauernhansl, T., Hompel, M. ten (Eds.), Handbuch Industrie 4.0 Bd.4. Springer, Berlin Heidelberg, pp. 1–31.

[2] Bayer, M., Bauer, F., 2023. Digitalisierung der Produktion im Mittelstand: Status quo und Handlungsempfehlungen für eine erfolgreiche Umsetzung, in: Schallmo, D.R.A., Lang, K., Werani, T., Krumay, B. (Eds.), Digitalisierung. Springer Fachmedien, Wiesbaden, pp. 469–485.

[3] Andrejić, M., Kilibarda, M., Pajić, V., 2018. A framework for assessing logistics costs, in: Quantitative models in economics. Publishing Centre, Belgrade, pp. 361–377.

[4] Bashatah, J.A., Elnaggar, G.R., 2025. Enhancing Warehouse Picking Efficiency Through Integrated Allocation and Routing Policies: A Case Study Towards Sustainable and Smart Warehousing. Applied Sciences 15 (20), 11186.

[5] Ruben, R.A., Jacobs, F.R., 1999. Batch Construction Heuristics and Storage Assignment Strategies for Walk/Ride and Pick Systems. Management Science 45 (4), 575–596.

[6] Hompel, M. ten, Henke, M., 2017. Logistik 4.0 – Ein Ausblick auf die Planung und das Management der zukünftigen Logistik vor dem Hintergrund der vierten industriellen Revolution, in: Vogel-Heuser, B., Bauernhansl, T., Hompel, M. ten (Eds.), Handbuch Industrie 4.0 Bd.4. Springer, Berlin Heidelberg, pp. 249–259.

[7] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–77.

[8] Schroth, T., Hummel, V., Leipzig, K. von, Schuhmacher, J., 2024. Development of a Simulation-Based Solution Concept for AI-Driven Clustering / Combination of Pick and Stow Operations to Improve Logistics Performance in SMEs, in: 2024 IEEE 20th International Conference on Automation Science and Engineering (CASE), pp. 1145–1150.

[9] Abdel-Basset, M., Abdel-Fatah, L., Sangaiah, A.K., 2018. Metaheuristic Algorithms: A Comprehensive Review, in: Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications. Elsevier, pp. 185–231.

[10] Elsawy, A., Warda, N., Ghoneimy, M., 2025. Improving Warehouse Efficiency with Dynamic Slotting Optimization Using Genetic Algorithm, in: 2025 International Conference on New Trends in Computing Sciences (ICTCS). IEEE, pp. 446–453.

[11] Gajsek, B., Dukic, G., Kovacic, M., Brezocnik, M., 2021. A Multi-Objective Genetic Algorithms Approach for Modelling of Order Picking. Int. j. simul. model. 20 (4), 719–729.

[12] Lee, I.G., Chung, S.H., Yoon, S.W., 2020. Two-stage storage assignment to minimize travel time and congestion for warehouse order picking operations. Computers & Industrial Engineering 139, 106129.

[13] Mirzaei, M., Zaerpour, N., Koster, R. de, 2021. The impact of integrated cluster-based storage allocation on parts-to-picker warehouse performance. Transportation Research Part E: Logistics and Transportation Review 146, 102207.

[14] Mojaver Tabrizi, A., Vahdani, B., Etebari, F., Amiri, M., 2023. A Three-Stage model for Clustering, Storage, and joint online order batching and picker routing Problems: Heuristic algorithms. Computers & Industrial Engineering 179, 109180.

[15] Peng, W., Guo, S., Li, H., 2021. Research on Picking Route Optimization of Multi-zone Warehouse Based on Traveling Salesman Problem and Simulated Annealing Algorithm. Journal of Physics Conference Series 1861 (1), 12031.

[16] Purwanto, C., Oktarina, R., 2024. Application of heuristic algorithms in warehouse order picking routes. JISS UNTIRTA 10 (2), 361.

[17] Chiang, T.-A., Che, Z.-H., Hung, C.-W., 2023. A K-Means Clustering and the Prim's Minimum Spanning Tree-Based Optimal Picking-List Consolidation and Assignment Methodology for Achieving the Sustainable Warehouse Operations. Sustainability 15 (4), 3544.

[18] Huynh, H.-D., Nguyen, D.D., Do, N.-H., Nguyen, K.-V., 2024. A Clustering Algorithm for Storage Location Assignment Problems in E-commerce Warehouses, in: Hai, N.T. et al. (Eds.), EAI International Conference on Renewable Energy and Sustainable Manufacturing. Springer Nature Switzerland, Cham, pp. 251–265.

[19] Xu, D., Tian, Y., 2015. A Comprehensive Survey of Clustering Algorithms. Ann. Data. Sci. 2 (2), 165–193.

[20] Zhuang, Y., Han, J., Hu, Y., 2020. Research on batching strategy of medical orders based on Canopy-K-means two-stage clustering algorithm, in: Proceedings of the 1st International Symposium on Artificial Intelligence in Medical Sciences, pp. 214–219.

[21] Schuhmacher, J., Hummel, V., Palm, D., Bauernhansl, T. Approach For Autonomous Control Of Intralogistics Considering Deterministic And Probabilistic Material Demand Information In Flexible Production Systems, in: Herberger, D., Hübner, M. (Eds.), Proceedings of the CPSL 2024.

[22] Zhou, H., 2024. Application of Internet of Things Technology in Intelligent Warehouse Management. ieim 7 (2).

[23] Munsberg, T.M., Hvam, L., 2021. Potential of Streamlining Warehouse Processes and IT integration to Increase Implementation of Automation, in: 2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE, pp. 46–50.

[24] Gabbrielli, A., Bordoy, J., Xiong, W., Fischer, G.K.J., Schaechtle, T., Wendeberg, J., Hoflinger, F., Schindelhauer, C., Rupitsch, S.J., 2023. RAILS: 3-D Real-Time Angle of Arrival Ultrasonic Indoor Localization System. IEEE Trans. Instrum. Meas. 72, 1–15.

[25] J, N.S.C., Wahab, N.H.A., Sunar, N., Ariffin, S.H.S., Wong, K.Y., Aun, Y., 2022. Indoor Positioning System: A Review. IJACSA 13 (6).

[26] ProGlove: Wireless Barcode Scanners for Modern Supply Chains, 2025. https://proglove.com/.

[27] Ergonomic glove for force measurement. https://www.innov8.eu/development/.

[28] Schroth, T., Schuhmacher, J., Hummel, V., Leipzig, K. von, 2025. Practical Validation of an Unsupervised Machine Learning Model for AI-Driven Cluster Analysis of Pick and Stow Operations within a Learning Factory, in: Louw, L., Hummel, V., Kock, I. de, Leipzig, K. von (Eds.), Advancing Learning Factories: Enabling Future-Ready Skills, vol. 1546. Springer Nature Switzerland, Cham, pp. 321–329.

[29] Universität Freiburg. Chair for Resilience Engineering - Prof. Dr. Alexander Stolz. Universität Freiburg. https://www.inatech.uni-freiburg.de/start_en/chairs/re?set_language=en.

[30] Telocate GmbH. https://en.telocate.de/.

Biography

Jan Schuhmacher (*1987) has been a research associate at Reutlingen University since 2014 and at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA since 2017. He is doctoral student at the Institute of Industrial Manufacturing and Management IFF at the University of Stuttgart since 2020. His major field of research is the design and optimization of changeable, autonomously controlled intralogistics systems.

Vera Hummel (*1962) has been a professor at Reutlingen University since 2010 and is initiator and head of the "Werk150 – The Factory of Reutlingen University". She is Dean of the Faculty NXT Sustainability and Technology of Reutlingen University and head of the teaching and research center for Value Added and Logistics Systems at Reutlingen University and the Reutlingen Center Industry 4.0, RZI 4.0 - a cooperation of the Fraunhofer Institutes IPA and IAO as well as Reutlingen University. She holds an extraordinary professorship at the Stellenbosch University, Industrial Engineering department.

8th Conference on Production Systems and Logistics

Artificial Intelligence-Driven Additive Manufacturing Use Cases: A Framework For Grounded Decisions

Konrad Schneidenbach, Carsten Feldmann
Münster University of Applied Sciences , Münster, Germany

Abstract

Purpose: This study aims to construct a decision framework for practitioners to improve understanding of industrial applications of artificial intelligence (AI) in additive manufacturing (AM) processes by systematically identifying and categorizing use cases. The study has two objectives: to analyze the current state of research and to highlight strategically relevant yet underexplored application areas.

Methodology: A systematic literature review of peer-reviewed journal articles and conference papers was conducted to identify the current state of research. Fourteen semi-structured expert interviews provided empirical data. The findings were synthesized using a two-dimensional heatmap combining AI functions (e.g. prediction, decision-making, reasoning) with AM process steps (e.g. design, material selection, post-processing).

Findings: The study presents the AI-AM decision framework as a conceptual model for analyzing the application of AI in AM. This constitutes a foundation for decisions. The statistically substantiated heat map reveals a high concentration of AI use cases for computer-aided design (CAD), computer-aided manufacturing (CAM) and quality assurance. This pattern is consistent regarding to the literature review and the expert interviews. However, literature more frequently indicates use cases addressing printing, while practitioners name the areas of calibration, ideation and requirements at a higher frequency. Core features of generative AI and self-operating systems, generating and acting AI functions occur solely in a few AM process stages. The heatmap reveals insufficiently examined intersections, especially in literature such as post-processing, calibration, printing and ideation related to specific AI functions. This study contributes by developing a conceptual heatmap and a decision framework of AI use cases in the AM process. Firstly, it advances the implementation of AI in AM by empirically identifying patterns substantiating practitioners in identifying potential use cases. These can enhance processes by automating tasks, managing risks, monitoring performance, and ensuring quality. Secondly, categorising emergent use cases via a heat map opens up valuable new directions for future research.

Keywords

Artificial Intelligence; AI; Additive Manufacturing; AM; Process Planning, Use Case Classification; Smart Manufacturing, Industry 4.0, Automation

1. Introduction

The expansion of Artificial Intelligence (AI) has profoundly impacted numerous aspects of processes. An important domain where AI could contribute to efficiency enhancements is expected in the manufacturing domain [1, 2, 3]. Manufacturing technologies that already include several digital process steps could embed

DOI: https://doi.org/10.15488/20985
ISSN: 2701-6277

AI in accordance with evolving technological dynamics. Additive manufacturing (AM) is such a representative of digital manufacturing technologies, due to its advancement and constant technologically innovations over the last decades [4, 5, 6]. The obtainable geometric freedom of the manufactured components, as well as numerous process parameters, results in high data volumes within this manufacturing method. Consequently, AI holds considerable potential to enhance the optimization of various parameters and their interactions within the manufacturing process, including data from Computer Aided Design (CAD), Computer Aided Manufacturing (CAM), materials, machine setup, and quality control. Due to this potential, there is growing interest in research into how AI can be applied in AM, and in particular which technological solutions are possible [7, 8, 9].

However, it is imperative to acknowledge the heterogeneity inherent in components manufactured using AM. This heterogeneity is influenced by various factors, including the most appropriate AM technology for a specific application, the type of material utilized, the software and interfaces employed, the technical requirements, the number of involved companies and the design of the supply chain, the prevailing standards, the category of the product, the geometry, and the quality specifications. This heterogeneity impedes the use of AI. In addition, AI itself provides specific functions that can be deployed to suit the specific manufacturing process step.

This heterogeneity complicates applying appropriate AI capability for use cases in the AM process. In a dynamic environment characterized by rapid advancements, a continual assessment and adaption of production processes is essential to retain competitive edge. This has not yet been extensively and systematically examined in the current literature: a research gap emerges concerning the development of a model-based and holistic perspective that relates AM processes to AI functions across various use cases [chapter 4]. Based on this research gap, the study addresses two research questions:

RQ1: *How can a decision framework be designed so that practitioners can analyse the use of AI in AM in a structured manner?*

RQ2: *Which combinations of AI functions and AM process stages are frequently encountered in both literature and practice?*

While researchers can utilize the identified gaps in subsequent studies, practitioners may apply the proposed model to evaluate use cases relevant to specific applications. This can provide guidance in the formulation of strategic decisions regarding where to invest, monitor developments, or refrain from focusing, based on market, research, and technical considerations. The present study offers two key contributions. Firstly, it presents a framework for identifying use cases for AI in the AM process, thereby supporting practitioners in their decision-making. Secondly, the categorisation of use cases via a heat map indicates relevance for new directions for future research.

2. Research methodology

To systematically assess the current state of the research, a **systematic literature review** was applied guided by the methodological approach established by Vom Brocke et al. [10]. To assure scientific robustness, the study includes peer-reviewed contributions and proceedings only. In the literature search, 33 keywords, like “3D printing” AND “artificial intelligence” AND “slicing”, were used in 145 combinations. To obtain comprehensive and holistic insights into the prevailing research focus areas, a search was conducted for various widely used keywords associated with technology or AM process stages. For instance, for the “design” AM process stage the keywords "CAD" and "3D modeling" were also applied. The list of keywords is available in [Appendix Nr. 1]. To guarantee rigor of the review, abstracts were sourced from principal peer-reviewed databases, including Web of Science, Emerald, IEEE Xplore and Sage. This initial search identified 14,943 results across all databases. To distil the genuinely appropriate literature, the PRISMA

methodology generated 47 relevant articles which deal with particular use cases aligned to thematic aspects and the framework which were included in the study, see [Appendix Nr. 2].

To complement the literature review with empirically grounded industry insights, 14 **expert interviews** were conducted. These were carried out both on-site and remotely via Zoom in November and December 2025. Following established methodological guidelines for expert interviewing [11, 12, 13], participants were selected through a targeted sampling based on specific knowledge, a substantial amount of professional experience over many years and organizational role. The sample included four CEOs as well as senior engineers, lead software developers, project and business developers, application managers, and researchers from renowned AM machine manufacturers, 3D-printing service providers, industrial user companies and AM software vendors.

All interviews were conducted as one-to-one conversations using an open, guideline-based format that balanced comparability with conversational flexibility. In line with best-practice recommendations [11, 12, 13], the interviewer employed non-suggestive prompts and situational probing to examine detailed information about the status quo, future and economic perspectives, implementation barriers and the range of use cases of AI implementation in the AM. Each interview lasted between 20 and 60 minutes, was audio-recorded with prior informed consent, and subsequently transcribed verbatim.

For the data analysis Kuckartz's qualitative content analysis was applied to structure content and to combine deductively derived main categories based on the research questions in the interview guide with inductively formed subcategories emerging from the material [14]. The software MAXQDA was used to support iterative coding and cross comparison. This methodological design ensures analytical rigor and provides a robust empirical foundation for understanding current AI adoption and integration, its challenges and future directions in AM.

3. Foundational Concepts for the AI-AM Heatmap

AI is empowering systems to mimic cognitive human behavior [15]. Human capabilities such as recognizing, interpreting, and deciding can now be recreated technically. The production of components using AM technologies is an expert domain which is requiring extensive know-how and experience by production operators [16]. Investigating AI-driven AM use cases can be conducted by mapping AI functions and AM processes within a two-dimensional heatmap. To structure the AI functions, the model from Hofmann et al. [15] is applied. Grounded in psychological insights into cognitive abilities they identify a task to AI function relation:

Perceiving: Capturing external data and transforming them into interpretable information.
Extracting: Uncovering and characterizing elements within raw data.
Reasoning: Deriving logical inferences about structural underlying relations.
Predicting: Anticipating forthcoming states or developments by estimating future outcomes.
Decision-Making: Deciding among existing discrete choices.
Generating: Composing new outputs based on learned patterns.
Acting: Conducting operations to achieve targeted objectives.

AM is specified as a process that utilizes 3D model data to produce components by joining material layer by layer additively in contrast to subtractive manufacturing [17]. This process involves a machine named 3D printer, material, digital data and machine control software, microelectronics and sensors as well as technicians and operators with extensive expertise to run this process. With regard to the tasks involved, the AM process can be conceptualized into the following stages:

1. **Ideation/Requirements:** Evaluating the optimal solution considering print technology, material, economical, technical and regulatory requirements.
2. **CAD:** Creating 3-dimensional technical drawing compliant with design guidelines.
3. **CAM:** Generating machine paths and defining material and machine processing parameters.
4. **Calibration:** Preparing the Printing-System.
5. **Printing:** Launching, monitoring and documenting the machining.
6. **Post-Processing:** Finishing the printed components, including strengthening and refining.
7. **Quality Assurance:** Inspecting and verifying the required quality parameters of the components.

The conjunction of AI functions with the aforementioned AM process stages yields a two-dimensional schema, the **AI-AM Heatmap**, see in chapter 4. Firstly, the intersections reveal zones in which the functions of AI contribute to the respective stage of the AM process. Secondly, the zones can be visually categorized into low, medium, and high concentration (cf. varying intensity of coloring in the zones) in accordance with the actual focus of research and/or practice, see chapter 4/5. Thirdly, a use case can be classified into the zones of the heatmap along the AM process, thus rendering the frequency visible.

4. State of the research and the AI-AM Heatmap

As illustrated in Table 1, the AI-AM Heatmap is presented in accordance with the research findings. This heatmap delineates the frequency zones that have been identified in the literature analysis.

A systematic analysis of AI use cases is paramount for deepening research in AM. AI and AM technologies are experiencing rapid technological advancement. Assessments guided by specific use cases provide substantial evidence of the application of AI functions within particular AM process stages, especially considering relevance in practical industrial environments. These analyses enable researchers to conduct a comprehensive evaluation of the performance, limitations, and integration challenges associated with AI implementation. By grounding research in documented applications rather than abstract assumptions, use case analysis enhances scientific rigor and ensures that insights into AI in AM remain relevant and transferable to industrial practice.

Current research on AI in AM shows clear patterns of high, medium, and low concentration across AI functions and AM process stages. **High-concentration** zones appear primarily in the zones CAD/generating, CAM/predicting, printing/predicting and in quality assurance/extracting and predicting. This is enabled by specific AI models for example topology optimisation, models for predicting structural behaviour, defect detection, sensor-based monitoring and data-driven quality control [18-45]. Especially metal printing technologies experience a lot of current thematic research focus in the stage of quality assurance, as well as medicine applications orthotics and lattice structures.

Medium-concentration zones include ideation and requirements/prediction and generating, CAM/ decision-making and generating, as well as printing/extraction. Literature presents in this zone isolated but promising approaches such as AI-based slicing optimisation, parameter tuning, sensor and databased monitoring of the printing process or distortion prediction [22, 46-58]. These contributions remain fragmented and often rely on case-specific datasets. **Low-concentration** zones occur at CAD/prediction and decision-making, calibration/predicting and acting, and post-processing/predicting. Only isolated studies about integrating AI into CAD systems or using this technology to enhance printer calibration and supporting the post-processing [59-63].

Table 1: AI-AM Heatmap: Literature analysis

At present, no research focus has been identified in the remaining "white" zones. Across all clusters, the literature mainly focus on isolated AM process stages/AI functions combinations, specific industries and machine technologies. While prediction and generating are frequent, holistic, cross-stage, end-to-end AI approaches remain largely unexplored. Existing approaches tend to optimise specific technical problems such as melt-pool dynamics or optimising mechanical properties for lattices [31, 49], rather than connecting these insights across the complete AM process stages. Only a few studies integrate multiple stages or combine AI functions into coherent decision frameworks [20, 21]. Moreover, most findings rely on non-standardised proprietary datasets, reproducibility, benchmarking, and cross-study synthesis [28, 52, 64].

Consequently, a clear **research gap** exists. A model that systematically links AM process stages with AI functions is absent in the literature. Such a model could assist to identify, classify, and evaluate AI-AM use cases in a structured manner. The AI-AM decision framework developed in this study directly addresses this gap by creating a conceptual foundation for bridging fragmented research streams, supporting empirical assessment, and guiding strategic decision-making.

5. Expert interview findings and the AI-AM Heatmap

To obtain a grounded perspective on practical application, expert interviews were conducted, see chapter 2, revealing the results presented in [Appendix Nr. 3]. The current implementation of AI is primarily still perceived as limited. Concurrently, the AI-AM integration is projected as having considerable future potential. The **prevailing AI applications** focus on image-based defect detection and process monitoring, data-analytical support for process parameters, and discrete AI functions in CAD/CAM, including generative design and automatic support generation. According to experts in the field, significant benefits are currently seen in the areas of quality assurance, print process monitoring, and automated print preparation. These areas can provide data availability, definable targets, and profit from direct efficiency gains.

Key barriers for AI integration include inadequacies in data quality and standardisation scarcity, limited interoperability between software and machine systems. Additionally, low confidence in the reliability of AI models occurs, especially in applications where certification and safety is fundamental. In addition, economic hurdles such as unclear responsibilities, data protection concerns and, in some cases, a low willingness to pay have been identified. Due to economic factors, such as proprietary data and production systems providing competitive advantages, responsibilities and liability within the supply chain. Data protection concerns are affecting the increasing implementation of AI in AM, the business model of hardware, software and material manufacturers, and service providers. Many experts see the ability to use data in an open model as a prerequisite for increasing AI use in the future. However, this also requires new approaches in business models that generate benefits for all business partners.

Forecasting the next five to ten years, experts expect AI to become a significantly larger importance alongside the AM process. An increase in automation and autonomous functions are expected. This includes self-optimizing slicing processes, CAD/CAM automation, AI-supported component orientation, adaptive process control based on sensory data, AI-supported post-processing steps, a predictable sintering process, and intelligent production planning for AM manufacturing.

6. The AI-AM Decision Framework for Identifying Use Cases

The literature review and expert interviews reveal that in conjunction with the implementation a high degree of complexity occurs by combining AI and AM technologies. This applies to the assessment and integration of this technology combination, as well as the calculation of prospective economic benefits. In parallel, both technologies are advancing at a rapid pace. Nevertheless, decisions are required about which new technological innovations to adopt and how to align these with existing technologies and processes. These

developments increase the necessity to maintain and strengthen competitive advantages. Regarding the first research question,

RQ1: *How can a decision framework be designed so that practitioners can analyse the use of AI in AM in a structured manner?*

this study proposes the following **AI-AM decision framework** to structure and handle the challenges identified, see [Appendix Nr. 4] including demo data. This framework consists of five elements: I. Action Plan; II. AI-AM Heatmap; III. Use Case Description & Analysis; IV. Use Case Navigator; and V. Decision Portfolio.

I. Action Plan provides a workflow for companies starting with a use case scouting; practitioners benefit from existing research and market perspectives by the AI-AM Heatmap (II.) and the Use Case Navigator (IV.) provided by this study. The next step is to describe, analyse and evaluate existing and potential use cases within the company in element III. The Use Case Navigator (IV.) can be used to verify the use case analysis in step S3 of the action plan or to utilise existing analyses as a blueprint. In this navigator, the AM process stages relevant to the respective use case are assigned to the respective AI function and coded.

For example, the zone C6 stands for the AM process stage 'CAM' (C) and the AI function "Generating" (6) for use case no. 1 'automation and optimisation of production data' if, for example, the AI edits G-code which is utilized as geometric code for instructing printer movements in the use case. The respective use case is subsequently classified in the AI-AM Heatmap (II). This Heatmap is a two-dimensional schema linking the AI Process stages with the AI Functions by visually indicating zones with a low, medium or high focus in research or practice, see chapter 2. The advantage of this approach is the immediate visibility of the dominant focus of research or practice within a specific zone of the use case. This enables the company to assess the extent to which the respective use case benefits from technological and market developments.

Furthermore, the company is capable of identifying zones where specific economic opportunities can be found due to the presence of minimal or non-existent competition. Moreover, this approach enables the comparison of individual use cases in terms of how promising they are. To add structure to this comparison, a decision portfolio can be created at the overall company level. In this portfolio, the use cases can be divided into the groups 'implement', "invest" and 'watch' gathered from the individual decision of the use case evaluation in III. This provides companies a systematic and consistent approach to direct their resources efficiently, unfold and transfer existing know-how and experience into new AI-AM technologies. Companies can define a percentage of each portfolio group and select use cases they identified and evaluated. For example, a company could apply a 10-10-80 percent rule; assign 10 percent each to group, implement, and invest by allocating the main resources for this thematic focus. A broader or shaper approach aside the pareto principle might suit for specific industries and companies. Overall, the AI-AM decision framework can support to handle complexity along rapid AI-AM development by improving decision quality to implement this technology combination to enhance competitiveness.

7. Discussion

In [Appendix Nr. 5], the synthesis of the literature-based heat map and the practitioner-oriented insights derived from the expert interviews is essential for obtaining a comprehensive understanding of AI adoption within the AM process.

While the literature analysis reflects the technological developments and research-driven perspectives on AI functions, the expert interviews reveal application-oriented priorities, operational constraints, and practical challenges occurring in the industrial plant. Integrating both viewpoints enables a more holistic assessment that incorporates technological maturity, practical feasibility, organisational requirements, and implementation barriers, as well as economic considerations. This combined perspective provides a robust empirical foundation from which a model can be developed that supports structured decision-making

regarding the deployment of AI across AM process stages. The resulting decision framework is presented in Chapter six.

The synthesis of the literature analysis and expert interviews uncovers a **high consistency of zones in the AI-AM heatmap.** These zones are with quality assurance/extracting and perceiving, CAM/predicting, CAD/generating and ideation_requirements/predicting and generating. The focus of research and practice in these zones is congruent. In contrast, practitioners are placing significantly more emphasis on all AI functions related to the AM process stages CAD and CAM.

In addition, the synthesis indicates that practitioners do not exclusively perceive significance and engagement in specific domains of individual zones, such as research. A more holistic and process-oriented perspective is adopted in regard to the implementation of artificial intelligence in asset management. This provides a foundation for novel research methodologies. A more promising approach, as suggested in Chapter 6, involves a shift in thinking to a more process-oriented framework across the individual AM process stages. This would entail establishing consistent links between upstream and downstream zones and thereby enhancing the overall efficiency and effectiveness of the AM process.

Regarding research question two, economic opportunities emerge in zones firstly where practitioners and research are extensively engaged. Companies can implement existing solutions and fine-tune technology to stay competitive for example in the AM process stage quality assurance, especially addressing the AI functions extracting and predicting. Secondly economic opportunities exist in zones and AM process stages where practitioners assume potentials with unsettled market supply. Optimising parameters using AI for post-processing procedures for example was named as an economically promising example by experts. Companies can invest in developing solutions to increase competitiveness for example in the CAD and CAM area including up- and downstream AM process stage. Thirdly economic opportunities could be discovered in untapped zones of the heatmap where a company is already in operation with their products and processes and therefore an economically profitable use case exists. Companies can invest in developing innovative solutions for their exclusive utilization or the creation of a novel product, to enhance competitiveness or revenues.

8. Conclusions and limitations

This study developed the AI-AM decision framework, for practitioners to derive decisions on AI use cases for AM in a systematic approach. It consists of an I. Action Plan as a Workflow guiding through the decision process, II. AI-AM heatmap combining AM Process Stages with AI Functions supported by extensive data from a literature review and expert interviews, III. Use Case Description and Analysis for deriving a detailed view on use cases, IV. Use Case Navigator approach to use existing use cases as a reference and V. Decision Portfolio for use case and categorical comparison. In regard to research question one, ***RQ1:*** *"How can a decision framework be designed so that practitioners can analyse the use of AI in AM in a structured manner?"*, the proposed framework supports practitioners in their decision-making.

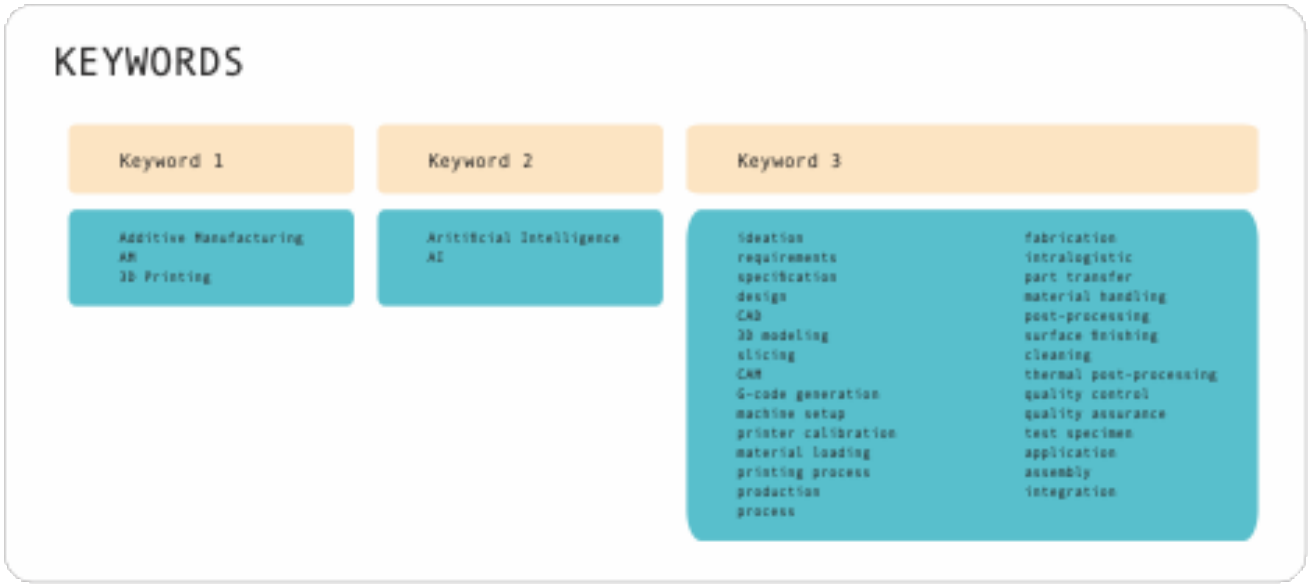
Regarding research question two, ***RQ2:*** *"Which combinations of AI functions and AM process stages are frequently encountered in both literature and practice?"*, the literature analysis reveals clear zones with high concentration in a few specific AI functions and AM process stage combinations, like CAD/generating, CAM/predicting, printing/predicting and in quality assurance/extracting and predicting as well as medium and low concentration zones presented in chapter four. The expert interviews uncover, that AI is already creating value in AM, but the current applications are still limited. Experts anticipate the highest benefit in AI-driven decisions on part-, print-technology-, material- selection regarding the defined requirements, automated production file generation with CAD/CAM-AI-integrations, data-driven process approaches with integration of large amounts of planning and manufacturing data, and AI-supported quality control.

However, this depends on developing data standards, business model incentives, interoperability, and trust in AI solutions.

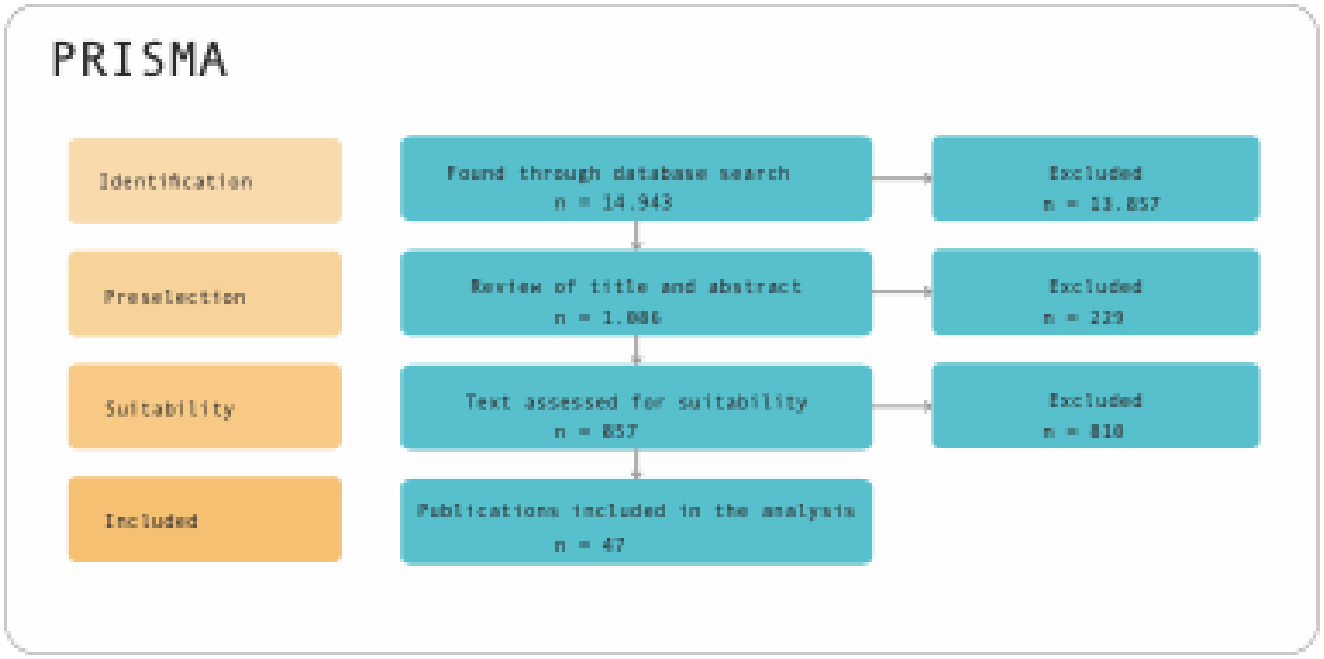
Current research lacks an integrated, model-driven approach that systematically connects AM process stages with AI functions to support a structured identification, classification, and evaluation of AI-enabled AM use cases. Economic opportunities emerge in zones of firstly high concentration of expert and research focus, secondly high practitioner focus and expected extensive economic potential and thirdly in untapped zones.

This study is subject to several **limitations**. Both research methods applied rely on human judgment. The systematic literature review depends on subjective assessments when determining the relevance of the included literature. Additionally, the rapid evolution of AI and AM technologies inherits, that consequently, emerging industrial applications, proprietary solutions, and unpublished implementations may not be fully captured. The empirical insights are based on a limited number of expert interviews. The sample size and integration of various industries restrict the generalizability of the findings. To substantiate the findings, subsequent investigation by experts and evaluation are therefore required. The framework does not provide quantitative validation of technical performance or economic impact. Its applicability therefore depends on context-specific factors such as data availability, organizational maturity, and strategic objectives. Subsequent research should validate the framework.

Appendix Nr. 1: Keywords

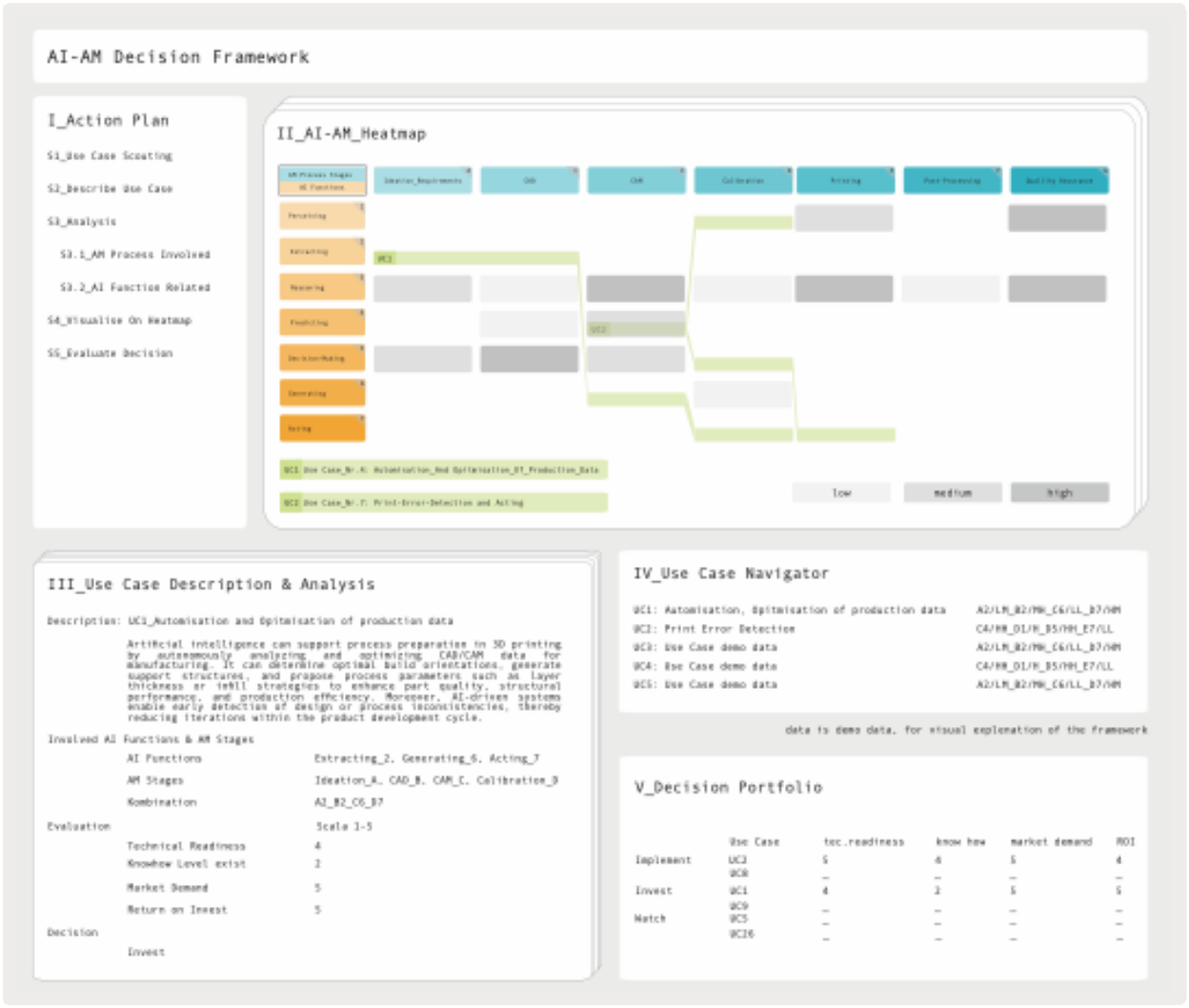


Appendix Nr. 2: PRISMA



Appendix Nr. 3: AI-AM Heatmap: Expert interviews

Appendix Nr. 4: AI-AM Decision Framework



Appendix Nr. 5: AI-AM Heatmap: Synthesis of literature analysis and expert interviews

References

[1] Mahajan, A., Singh, G., Devgan, S., Singh, G., Kaur, P., 2026. Next-generation manufacturing: leveraging AI for industrial innovation and growth. Computers & Industrial Engineering 211, 111618.

[2] Gao, R.X., Krüger, J., Merklein, M., Möhring, H.C., Váncza, J., 2024. Artificial Intelligence in manufacturing: State of the art, perspectives, and future directions. CIRP Annals 73(2), 723-749.

[3] Arinez, J.F., Chang, Q., Gao, R.X., Xu, C., Zhang, J., 2020. Artificial intelligence in advanced manufacturing: current status and future outlook. Journal of Manufacturing Science and Engineering 142(11), 110804.

[4] Zhou, L., Miller, J., Vezza, J., Mayster, M., Raffay, M., Justice, Q., Al Tamimi, Z., Hansotte, G., Sunkara, L.D. and Bernat, J., 2024. Additive manufacturing: a comprehensive review. Sensors 24(9), 2668.

[5] Gibson, I., Rosen, D., Stucker, B., Khorasani, M., Rosen, D., Stucker, B., Khorasani, M., 2021. Additive manufacturing technologies, 3rd ed. Springer Nature Switzerland AG, Cham.

[6] Adapa, V.S.K., Kalidindi, S.R. and Saldana, C.J., 2025. Rapid Development of Metal Additive Manufacturing Using Artificial Intelligence/Machine Learning and High-Throughput Material Testing. Annual Review of Materials Research 55, 175-201.

[7] Tamir, T. S., Hua, X., Jiang, J., Leng, J., Xiong, G., Shen, Z., Liu, Q., 2025. Data-driven and physics-assisted machine learning approach for warpage classification and process parameter optimization in a 3-D-printed beltclip. IEEE Transactions on Computational Social Systems 12 (4), 1637–1651.

[8] Khod, S., Goswami, M., Dvivedi, A., 2025. Additive manufacturing processes protocol prediction by Artificial Intelligence using X-ray Computed Tomography data. Progress in Additive Manufacturing 10 (9), 6691–6709.

[9] Goh, G. D., Sing, S. L., Yeong, W. Y., 2021. A review on machine learning in 3D printing: applications, potential, and challenges. Artificial Intelligence Review 54(1), 63-94.

[10] vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A. 2009. Reconstructing the Giant: on the Importance of Rigour in Documenting the Literature Search Process. European Conference on Information Systems 17, 2206-2217.

[11] Mieg, H.A., Näf, M., 2005. Experteninterviews in den Umwelt- und Planungswissenschaften. Eine Einführung und Anleitung. 2. Aufl. Institut für Mensch-Umwelt-Systeme (HES), ETH Zürich.

[12] Kurzrock, B.-M., 2014. Anleitung für Experteninterviews im Rahmen wissenschaftlicher Arbeiten am Fachgebiet Immobilienökonomie. Technische Universität Kaiserslautern.

[13] Taherdoost, H., 2022. How to Conduct an Effective Interview; A Guide to Interview Design in Research Study Authors. International Journal of Academic Research in Management 11 (1), 39–51.

[14] Kuckartz, U., Rädiker, S., 2024. Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und künstlicher Intelligenz, 6. Auflage. Beltz Juventa, Weinheim.

[15] Hofmann, P., Jöhnk, J., Protschky, D., Urbach, N., 2020. Developing purposeful AI use cases: a structured method and its application in project management. Proceedings of the 15th International Conference on Wirtschaftsinformatik (WI) 15, 33–49.

[16] Schneidenbach, K., Feldmann, C., 2024. Impact of additive manufacturing and parametric design on the structure and economic efficiency of construction supply chains. IFAC PapersOnLine 58(19), 1000–1005.

[17] ASTM, 2021. ISO/ASTM 52900:2021(en), additive manufacturing — general principles — fundamentals and vocabulary. Available at: https://www.iso.org/obp/ui/#iso:std:iso-astm:52900:ed-2:v1:en (accessed 12 September 2025).

[18] Mohseni, M., Khodaygan, S., 2024. Design for additive manufacturing of topology-optimized structures based on deep learning and transfer learning. Rapid Prototyping Journal 30 (7), 1411–1433.

[19] Adapa, V.S.K., Kalidindi, S.R., Saldana, C.J., 2025. Rapid development of metal additive manufacturing using artificial intelligence/machine learning and high-throughput material testing. Annual Review of Materials Research 55, 175–201.

[20] Ghimire, R., Raji, A., 2024. Use of artificial intelligence in design, development, additive manufacturing, and certification of multifunctional composites for aircraft, drones, and spacecraft. Applied Sciences-Basel 14 (3).

[21] Khod, S., Goswami, M., Dvivedi, A., 2025. Additive manufacturing processes protocol prediction by artificial intelligence using X-ray computed tomography data. Progress in Additive Manufacturing 10 (9), 6691–6709.

[22] Duan, D.L., Wu, W.Z., 2025. Inverse design of lattice structures with target mechanical performance via generative adversarial networks considering the effect of process parameters. Advanced Engineering Informatics 65 (B).

[23] Tsegay, F., Ghannam, R., Daniel, N., Butt, H., 2023. 3D printing smart eyeglass frames: a review. ACS Applied Engineering Materials 1 (4), 1142–1163.

[24] Yu, H.Z., Li, W., Li, D., Wang, L.J., Wang, Y., 2024. Enhancing additive manufacturing with computer vision: a comprehensive review. International Journal of Advanced Manufacturing Technology 132 (11-12), 5211–5229.

[25] Tamir, T.S., Hua, X., Jiang, J., Leng, J., Xiong, G., Shen, Z., Liu, Q., 2025. Data-driven and physics-assisted machine learning approach for warpage classification and process parameter optimization in a 3D-printed BeltClip. IEEE Transactions on Computational Social Systems 12 (4), 1637–1652.

[26] Gomez-Omella, M., Flores, J., Sierra, B., Ferreiro, S., Hascoet, N., Chinesta, F., 2023. Optimizing porosity detection in wire laser metal deposition processes through data-driven AI classification techniques. Engineering Failure Analysis 152.

[27] Li, Q., Huang, T., Liu, J., Wang, S., 2025. Causality-guided fault diagnosis under visual interference in fused deposition modeling. Engineering Applications of Artificial Intelligence 143.

[28] Zhang, H., Yu, Y., Liu, X., Liang, N., Kim, Y., Shin, D., Moon, S., Choi, J.P., 2025. Autonomous printing process optimisation and in-situ anomaly detection in fused deposition modelling using an integrated data-driven approach. Virtual and Physical Prototyping 20 (1).

[29] Bakas, G., Bei, K., Skaltsas, I., Gkartzou, E., Tsiokou, V., Papatheodorou, A., Karatza, A., Koumoulos, E.P., 2022. Object detection: custom trained models for quality monitoring of fused filament fabrication process. Processes 10 (10).

[30] Kuriachen, B., Jeyarai, R., Raphael, D., Ashok, P., Sundari, P.S., Paul, A., 2025. Defect detection in fused deposition modelling using lightweight convolutional neural networks. Engineering Applications of Artificial Intelligence 141.

[31] Shang, X., Talbot, A., Li, E., Wen, H., Lyu, T., Zhang, J., Zou, Y., 2025. Accurate inverse process optimization framework in laser directed energy deposition. Additive Manufacturing 102.

[32] Fang, Q., Xiong, G., Zhou, M.C., Tamir, T.S., Yan, C.B., Wu, H., Shen, Z., Wang, F.Y., 2024. Process monitoring, diagnosis and control of additive manufacturing. IEEE Transactions on Automation Science and Engineering 21 (1), 1041–1067.

[33] Muhammad, W., Kang, J., Ibragimova, O., Inal, K., 2023. Experimental investigation and development of a deep learning framework to predict process-induced surface roughness in additively manufactured aluminum alloys. Welding in The World 67 (4), 897–921.

[34] Ulkir, O., Bayraklilar, M.S., Kuncan, M., 2024. Raster angle prediction of additive manufacturing process using machine learning algorithm. Applied Sciences-Basel.

[35] Abhilash, P.M., Ahmed, A., 2023. Convolutional neural network-based classification for improving the surface quality of metal additive manufactured components. The International Journal of Advanced Manufacturing Technology 126 (9-10), 3873–3885.

[36] Asadi, R., Queguineur, A., Wiikinkoski, O., Mokhrarian, H., Aihkisalo, T., Revuelta, A., Ituarte, I.F., 2024. Process monitoring by deep neural networks in directed energy deposition: CNN-based detection, segmentation, and statistical analysis of melt pools. Robotics and Computer-Integrated Manufacturing 87.

[37] Bordekar, H., Cersullo, N., Brysch, M., Phillip, J., Hühne, C., 2025. Explainable artificial intelligence for automatic defect detection in additively manufactured parts using CT scan analysis. Journal of Intelligent Manufacturing 36 (2), 957–974.

[38] Piran, F.J., Poduval, P.P., Barkam, H.E., Imani, M., Imani, F., 2025. Explainable differential privacy-hyperdimensional computing for balancing privacy and transparency in additive manufacturing monitoring. Engineering Applications of Artificial Intelligence 147.

[39] Mottafegh, A., Rohimmahtunnissa, A., Kim, D.-P., 2025. Rapid prototyping of 3D-printed microchannel heat exchangers via AI-driven design and optimization. Chemical Engineering Journal 520.

[40] Wang, X., Li, Y., Liang, Z., Du, R., Song, T., 2025. High-fidelity, personalized cardiac modeling via AI-driven 3D reconstruction and embedded silicone rubber printing. Experimental Biology and Medicine 250.

[41] Abinaya, A., Veeramakali, T., 2025. AI-powered 3D printing error detection and optimization system. 2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE).

[42] Vaid, H., Shubham, P., Sharma, P., Vajpeyi, R., 2021. Prediction of printing parameters for minimal dimensional variation in 3D printed parts by using artificial intelligence. 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN).

[43] Kumar, D., Liu, Y., Song, H., Namilae, S., 2023. Explainable deep neural network for in-plain defect detection during additive manufacturing. Rapid Prototyping Journal (2024) 30 (1), 49–59.

[44] Pan, D., Jia, F., Zhou, M., Liu, H., Yan, J., Zhu, L., Yang, M., Jiang, Z., 2025. Advances in 3D printing: microfabrication techniques and forming applications. Micromachines 16 (8).

[45] Iuso, D., Chatterjee, S., Cornelissen, S., Verhees, D., Beenhouwer, J., Sijbers, J., 2024. Voxel-wise segmentation for porosity investigation of additive manufactured parts with 3D unsupervised and (deeply) supervised neural networks. Applied Intelligence 54 (24), 13160–13177.

[46] Rojek, I., Mikołajewski, D., Dostatni, E., Macko, M., 2020. AI-optimized technological aspects of the material used in 3D printing processes for selected medical applications. Materials 13 (23).

[47] Kanagalingam, S., Dalton, C., Champneys, P., Boutefnouchet, T., Fernandez-Vicente, M., Shepard, D.E.T., Wimpenny, D., Thomas-Seale, L.E.J., 2023. Detailed design for additive manufacturing and post processing of generatively designed high tibial osteotomy fixation plates. Progress in Additive Manufacturing 8 (3), 409-426.

[48] Zhang, L., Bibi, F., Hussain, I., Sultan, M., Arshad, A., Hasnain, S., Alarifi, I.M., Alamir, M.A., Sajjad, U., 2023. Evaluating the stress-strain relationship of the additively manufactured lattice structures. Micromachines14 (1).

[49] Ali, M., Sajjad, U., Hussain, I., Abbas, N., Ali, H.M., Yan, WM., Wang, CC., 2022. On the assessment of the mechanical properties of additively manufactured lattice structures. Engineering Analysis with Boundary Elements 142, 93–116.

[50] Logeshwaran, A., Elsen, R., Nayak, S., 2024. Artificial intelligence-based 3D printing strategies for bone scaffold fabrication and its application in preclinical and clinical investigations. ACS Biomaterials Science & Engineering 10 (2), 677–696.

[51] Ozalp, A., Aldemir, A., 2025. Artificial intelligence-based displacement capacity prediction tool for three-dimensional printed concrete walls. Structural Concrete.

[52] Paraskevoudis, K., Karayannis, P., Koumoulos, E.P., 2020. Real-time 3D printing remote defect detection (stringing) with computer vision and artificial intelligence. Processes 8 (11).

[53] Barjuei, E.S., Capitanelli, A., Bertolucci, R., Courteille, E., Mastrogiovanni, F., Maratea, M., 2024. Digital workflow for printability checking and prefabrication in robotic construction 3D printing based on artificial intelligence planning. Engineering Applications of Artificial Intelligence 133 (C).

[54] Thompson, M., Sarvestani, H.Y., Sohrabi-Kashani, A., Kiyani, E., Filippi, E., van Egmond, D.A., Rahmat, M., Ashrafi, B., Karttunen, M., 2025. Machine learning-optimized stochastic Voronoi lattices for enhanced mechanical performance. Engineering Applications of Artificial Intelligence 160 (C).

[55] Papacharalampopolous, A., Stavropoulos, P., 2025. Lighting design for industry 5.0 and the New Bauhaus: additive manufacturing of luminaires. 2025 Tenth Conference on Lighting (Lighting).

[56] Haque, S., Tanveer, M.H., Voicu, R.C., 2025. AI-driven 3D printing: generative G-code for quality and efficiency. SoutheastCon 2025.

[57] Khod, S., Goswami, M., Dvivedi, A., 2025. Additive manufacturing processes protocol prediction by artificial intelligence using X-ray computed tomography data. Progress in Additive Manufacturing 10 (9), 6691–6709.

[58] Shen, W., Zhang, P., Li, W., Qin, H., 2025. Additive manufacturing for functional design: a review of capabilities, strategies and applications. Rapid Prototyping Journal 1–27.

[59] Nasr, M.M., Anwar, S., 2025. Developing an intelligent approach based on ANFIS and advanced NSGA-III for improving the turning performance of additively manufactured γ-TiAl alloy. International Journal of Advanced Manufacturing Technology 137 (7–8), 3333–3358.

[60] Boretti, A., Huang, AJ., 2024. AI-driven DfAM of aeronautical hydrogen gas turbine combustors. International Journal of Hydrogen Energy 77, 851–862.

[61] Behera, R., Chan, TC., 2025. Dynamic analysis of 3D printer through simulation, experimental validation, and deep learning-based prediction. Journal of The Chinese Institute of Engineers.

[62] Leong, W.Y., 2025. Structural optimization and performance evaluation of Spain's 3D-printed concrete bridges using AI-based design approaches. 2025 10th International Conference on Applying New Technology in Green Buildings (ATiGB).

[63] Kelly, D., Sergis, V., Blanco, L., Mason, K., Daly, A.C., 2025. Autonomous control of extrusion bioprinting using convolutional neural networks. Advanced Functional Materials 35 (30).

[64] Riffel, K.C., Bawa, R., Chan, J., Adhami, R., Gofman, D., Raimrez, A.J., 2025. Artificial intelligence and statistical mapping applied to additive manufacturing toolpath optimization in Wire-Arc DED. Integrating Materials And Manufacturing Innovation 14 (3), 425–441.

Biography

Dipl.-Kfm. Konrad Schneidenbach is a lecturer in Digital Transformation and Digital Production Processes. He is an expert in 3D printing, digital platforms and planning automation. At Münster University of Applied Sciences, he is the head of the digi.lab at the Faculty of Business. As Project Manager at the Institute for Process Management and Digital Transformation (IPD), he supports startups and established companies to develop product-, process- and business model innovations.

Prof. Dr. Carsten Feldmann is a professor of business process management at the University of Applied Sciences Münster. He is the chairman of the Institute for Process Management & Digital Transformation (IPD). Before he worked in various leadership positions at Bosch, e.g. as a plant manager and Director of Worldwide Manufacturing Coordination, and Accenture. His research focuses on process automation, digital transformation, and supply chain management.

8th Conference on Production Systems and Logistics

A Conceptual Framework for Evaluating the Transferability of Production Management Decisions to AI

Joshua Kaiser, Carl René Sauer, Norman Müller, Luisa Stracke, Peter Burggräf

Department of Mechanical Engineering, Chair of International Production Engineering and Management (IPEM, University of Siegen, Siegen,Germany)

Abstract

The field of artificial intelligence (AI) has advanced rapidly in recent years, demonstrating remarkable capabilities in solving complex decision-making tasks once considered exclusive to human intelligence. The potential of AI has been showcased through its success in strategic games, where AI systems have already surpassed human performance in structured, rule-based environments. This achievement raises the fundamental question of whether similar decision-making abilities can be applied to production management (PM) context that involves uncertainty, dynamic interactions, and strategic trade-offs. However, in the field of PM, the systematic application of AI for decision-making is still emerging and no comprehensive and controlled approach has been established to systematically compare human and AI performance. The research project "Manager ex Machina – Investigation of the Transferability of Production Management Decisions to AI Systems" aims to develop a simulation-based method that enables this evaluation and comparison of decision-making performance. Therefore, a simulation will be developed as a serious game that enables both human participants and AI agents to make PM decisions within a controlled and reproducible environment. The decision scenarios forming the serious game will be derived from expert interviews with specialists in PM and AI to identify decisions with the highest automation potential. The resulting data from the serious game will then be analyzed to evaluate the effectiveness and limitations of AI performance in these specific PM decisions. This allows to establish a conceptual framework for evaluating AI decision-making performance that can be transferable to other managerial domains, including workforce planning and supply-chain management. Future studies should focus on validating and refining the proposed concept through empirical studies and the development of a serious game. This will contribute to the development of a systematic benchmarking framework for evaluating human – AI performance in decision-making contexts.

Keywords

Human-AI collaboration; Simulation-based evaluation; Serious game; hybrid intelligence

1. Introduction

Digital transformation and the increasing use of artificial intelligence (AI) are fundamentally changing industrial value creation systems and production management (PM), as production environments become increasingly data-driven and interconnected [1]. Technological developments, such as increasing computational power and the miniaturization of embedded systems are enabling the deployment of cyber-physical systems as autonomous and responsive agents capable of learning from large datasets and automating routine tasks [2–4]. These technological developments lead to a growing tension between

DOI: https://doi.org/10.15488/20986
ISSN: 2701-6277

emerging autonomous decision-making capabilities and traditional forms of human managerial decision-making [5]. At the same time, modern production environments impose an increasing number of requirements and constraints, making it progressively more difficult for human decision-makers to oversee all relevant information and respond effectively. This growing complexity creates additional momentum for exploring more autonomous technological capabilities [6]. At the operational level, data-driven, real-time decision support is already established in many industrial applications, as reflected in the growing use of predictive maintenance systems [7]. In PM, however, comparable AI-based support remains uncommon, especially for strategic and tactical decision-making, even though increasing AI adoption is widely expected and the need for integration is apparent [9,8]. Existing approaches in PM predominantly address narrowly defined operational optimization tasks, such as capacity or scheduling [7, 13], while implications for more complex tactical and strategic decisions remain largely unexplored [10, 11]. However, these tactical and strategic decisions particularly describe a large part of PM [12]. In general, companies tend to proceed with caution in making experimental changes to tactical and strategic decisions or decision-making processes, unless there is sufficient validation of the expected performance [10]. Long-term, strategic decisions only become evident after extended periods, and AI-performance evaluation in ongoing operations can have operational risks and therefore lead to potential productivity losses [10]. Consequently, a benchmark is required to assess AI decision-making performance in advance. Thus, this paper presents a concept for a simulation-based human-AI benchmark within the PM domain. The proposed approach establishes a structured foundation for the consistent and comparable evaluation of AI performance in tactical and strategic, practice-oriented decision-making scenarios. This paper is structured as follows: Section 2 presents a summary of the current state of the art, Section 3 details the solutional approach to develop the assessment of human-AI performance in PM decisions, and Section 4 reports the preliminary findings. The final section of the paper discusses implications and outlines future research directions.

2. State of the art

The core definition of PM is to control industrial value-creation processes, guaranteeing the efficient transformation of resources into marketable goods in accordance with quality and organizational requirements [13]. Another objective is to sustain or potentially increase the economic value generated through the comprehensive utilization of available resources [14]. Therefore, PM encompasses all the decisions required for the management of industrial value-creation processes and many sub-decisions are highly interdependent and influence each other over extended periods [12,15]. To further structure decision-making in PM, decisions are commonly differentiated into operational, tactical, and strategic levels [16,17]. Operational decisions govern short-term project execution and ensure the timely implementation of planned work packages [18]. Tactical decisions translate strategic intent into mid-term planning and control, for example by shaping resource allocations and phase plans [19]. Finally, strategic decision-making concerns the project long-term orientation and justification, including its prioritization and alignment with organizational objectives [18].

AI, in addition, refers to computer-based systems designed to emulate human cognitive capabilities by learning from data and making autonomous decisions in ways that resemble human reasoning [20]. These systems are able to acquire knowledge from experience, adapt to new circumstances, and generate solutions for decisions that typically require human intelligence [21]. These systems are capable of identifying patterns in changing data and improving their performance autonomously without being explicitly programmed for every possible situation [23,22].

Scientific discourse traditionally differentiates between the respective strengths of autonomous systems and human decision-makers. Autonomous systems are expected to outperform traditional systems in clearly defined and structured decisions, where they can operate with high speed. In contrast, humans are generally regarded as superior in decisions that require inductive reasoning and flexible adaptation to incomplete

information [24–26]. Activities involving creativity or social judgment are considered to remain within the domain of the human decision maker in the long term [27]. Regarding PM this becomes evident, as applications in PM remain mostly limited to well-structured operational decisions like process planning while the maturity of ML is growing in areas such as demand forecasting or automated visual inspection [10,28,29]. More complex or long-term decisions like the handling of unplanned disruptions (e. g., machine failures or worker absences) are rarely addressed in existing research. Consequently, the performance of AI within these tactical and strategic decisions in PM needs to be systematically assessed [11].

Benchmarks such as GLUE (General Language Understanding Evaluation) predominantly measure language understanding or numerical problem-solving in isolated and short-horizon decisions. These benchmarks fail to capture the abilities that are essential for managing or planning decisions in real systems, such as long-term coherence and adaptability to dynamic changes [30]. An alternative and well-established approach for evaluating human and AI performance can be found in the domain of games. Strategic games such as chess or Go have served as AI-benchmarks for decades since they offer defined rules, well-deterministic state spaces, and objective performance metrics [31]. In these environments, performance can be objectively quantified through metrics such as Elo ratings or by comparing against human reference groups [31]. Modern reinforcement learning (RL) approaches, including Deep Q-Networks have achieved superhuman performance in these structured game settings [33,32]. In chess, for instance, AlphaZero outperforms all existing traditional chess programs after only a few hours of training [31].

While the presented benchmarks illustrate that AI can achieve exceptional performance on narrowly defined, highly formalized decisions, their applicability to complex domains such as PM with long-term dependencies and incomplete information is limited [34]. This does not imply that benchmark principles derived from deterministic game environments such as chess or Go are fundamentally non-transferable to stochastic PM settings. Even in chess, fixed rules and full observability do not eliminate strategic uncertainty. The opponent's response and the downstream consequences of a move remain unknown at decision time. Similarly, PM decisions are based on an assessment of the current system state under information that is only partially available [35]. At the same time, PM typically involves a larger set of interacting decision variables and is subject to external disturbances, which increases the complexity of both state estimation and planning [36]. Nevertheless, game-based environments provide controlled and repeatable conditions for evaluation, whereas a direct benchmark in real production systems is hardly feasible because complex strategic or tactical decisions often unfold their effects only after months or years. In addition, test phases carry operational risks and potential productivity losses, complicating the execution of experiments [37,38]. Emerging research therefore aims to leverage more complex game environments as benchmarks, such as business simulation games with multi-month cycles and multi-day decision-making that represent extended time horizons and multiple interaction steps [40,39]. Although these approaches demonstrate the potential of game-based environments for human-AI performance comparisons, they still lack a consistent scientific foundation. They rarely include a systematic modeling of the underlying decision situations, a clear transferability of results to real management contexts or a precise representation of complexity and uncertainty. Moreover, the performance of human and artificial decision-makers is not assessed through a unified benchmark significantly limiting the comparability of results. In addition, a comparison of performance in PM should reflect that decision problems vary in difficulty. Consequently, the best-performing party is not necessarily the one that solves the highest number of tasks, but the one that achieves the highest overall decision quality, including on the most difficult decision situations. This requirement can be addressed by linking the benchmark to Item Response Theory (IRT), which enables difficulty-adjusted performance assessment rather than relying on unweighted success counts.[41].

This presented research results are indicating that current AI benchmarks lack the capacity to methodically capture the complex and time-dependent decision-making processes that are characteristic for PM. This gap highlights the need for a concept that enables controlled and reproducible comparisons between human and

AI decision-makers without interfering with actual business operations. From these findings, the guiding research questions emerges: “How can a standardized, controllable and repeatable methodology be developed to objectively compare human and AI decision-making in production management?” The initial observation that game environments are used to compare the performance of AI and humans, and the fact that AI has already surpassed human capabilities in structured game settings, serves as the foundation for this concept. The underlying hypothesis is that transferring this benchmarking logic to production systems enables a direct comparison of human and AI decision-making within a controlled and reproducible environment. To examine this assumption, a simulation-based benchmark concept with several steps is being developed, in which both an AI agent and human participants work through identical decision problems. The simulation-based approach provides an abstracted and realistic representation of a production system, allowing performance differences to be measured objectively through defined key performance indicators (KPIs). This approach is presented within a suitable conceptual benchmark in Chapter 3.

3. Solutional approach

Following the challenges outlined in the previous chapter, a methodological approach is being developed to assess the performance of AI in PM for decisions traditionally carried out by human. The approach must ensure that empirical, methodological, and practical requirements for comparing human and AI decision-making performance are coherently addressed and systematically validated. The three main objectives are:

- Objective A: Identification and structuring of relevant production-management decisions and their translation into an abstracted, coherent simulation environment that reflects the essential system logic essential system logic, information flows, and decision structures
- Objective B: Development and integration of suitable AI agents into the simulation, followed by a controlled empirical study in which human participants and AI systems solve identical decision tasks under comparable conditions
- Objective C: Evaluation and comparison of human and AI performance using a unified assessment model and validation of the simulation results through expert feedback and their transfer into selected real-world production environments

To realize these objectives six corresponding work phases that form the methodological backbone of the development are suggested (cf. figure 1).

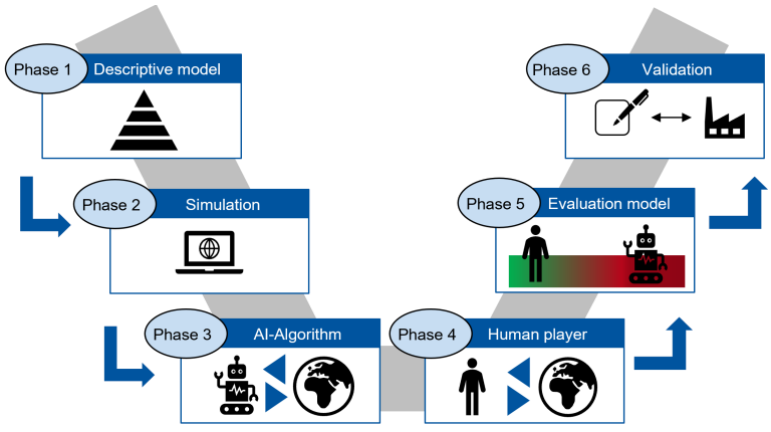


Figure 1: Work packages to develop a human-AI benchmark in a V-Model

Phase 1: Descriptive model for a production system

The first step in the methodological development is to determine which types of decisions in PM should be considered for the development of a simulation-based comparison between human and artificial decision-makers. PM is characterized by a multitude of heterogeneous decisions that differentiate significantly in

temporal horizon, available information, involved actors, and structural complexity [42]. To develop a realistic decision-making space, a systematic approach is essential for identifying and structuring these decisions. Phase 1 therefore follows a conceptual approach based on an exploratory analysis of industrial decision-making processes. As outlined in the study by Aghamanoukjan (2009) [43], the procedure could involve qualitative interviews with production managers or the analysis of standard planning and control processes. The objective is to gain a comprehensive understanding of the decision situations encountered in practice, the information available to decision-makers and the typical steps of the respective decision processes. The insights gained are subsequently processed through a four-step qualitative content analysis following Mayring (2014) [44] and are integrated into a structured description model. This model captures essential elements of each considered decision. At the same time, the model must be sufficiently abstract to be transferred into a simulation or game-based environment without losing its connection to real production practice. This requires simplifying real decision processes while preserving their core characteristics. The challenge is to reduce the complexity of production decisions to a level understandable for humans and computable for AI, while maintaining the essence of the decision situation. It includes establishing clear decision criteria and implementing a consistent research design that enables objective performance comparison. While any decision could theoretically be selected and passed through phases 2-6, the structured approach in phase 1 ensures that the benchmark begins with a decision where the methodological effort is aligned with its best cost-benefit value and has potential to generate relevant insights for the subsequent development steps. Phase 1 thus establishes the methodological foundation for all subsequent phases.

Phase 2: Definition of the simulation environment

Phase 2 is based on the findings from phase 1, focusing on converting the structure into a simulation that allows for collaborative decision-making between humans and AI under identical conditions. The objective is not to replicate the complete PM, but rather to establish a controlled experimental environment that simplifies essential characteristics of real decisions into a game-like format. The first step is to transform decisions from phase 1 into a reduced and consistent representation of industrial system logic. To achieve a balance between human comprehensibility and AI tractability, the simulation abstracts the essential logic of production systems, integrating aspects such as resource availability and the handling of unexpected events into a coherent and simplified representation. The key challenge is to reconcile realism and manageability. The simulation is required to be realistic enough to reveal meaningful performance differences, as well as structured to function as a scientific experiment. Because no established scientific methodology exists for systematically designing a serious game for human-AI performance comparison in PM, phase 2 develops a conceptual framework. This framework defines the elements necessary for such an environment to operate both as a game and as a research instrument. It specifies core game mechanics, Interfaces to the AI-Algorithm and defines the way how users interact with the environment. It also outlines the narrative structure, ensuring that decision scenarios are understandable to human players. In addition, the framework defines the analytical components required for a scientifically robust comparison, including a unified objective and KPI structure, consistent decision rules and mechanisms ensuring that humans and AI truly operate under comparable conditions. Based on these requirements, the core game mechanics can be organized into a tutorial phase with a stepwise introduction of game elements, followed by a round-based practice phase and a subsequent assessment phase where acquired knowledge is applied under standardized conditions. This staged design familiarizes participants with the interaction logic and reduces the risk that performance differences reflect video-game proficiency rather than PM-relevant decision competence, ensuring that the assessment does not primarily reflect gaming literacy and thus limiting gamification bias. Phase 2 establishes the conceptual foundation, integrating the modeling of the decision space with empirical investigation.

Phase 3: AI-Algorithm implementation and training

Following the development of the simulation environment, phase 3 focuses on selecting, developing and training an AI-Algorithm capable of operating within the same environment as human participants. The

objective is to define the requirements for an AI model that can represent the decisions identified in phase 1 and execute them within the simulation. Ensuring comparability requires that humans and AI have access to identical information, action options, and feedback. This leads to the question of how an artificial system must be structured to handle complex PM decisions sequentially and in a transparent manner. To ensure effective communication with an AI model, the simulation state information and available actions must be formatted in a clear and comprehensible structure. At the same time, the interaction between the AI system and the environment must follow the same principles and informational constraints as those applying to human players. When selecting an appropriate AI-Algorithm for performance comparisons in PM, it is essential to ensure that the selected AI is capable of handling the complex decision types characteristic of PM, in particular tactical and strategic decisions. RL represents a promising approach for designing such a system. It is particularly suitable for PM decisions because it learns decision strategies through repeated interaction with an environment, regardless of the availability of historical datasets [45]. This is especially advantageous in PM, where reliable data for long-term or strategic decisions are often limited. An RL agent can improve its decision-making behavior exclusively through feedback from the simulated environment, enabling it to discover complex relationships and iteratively develop appropriate strategies.

Formally, each decision problem is represented as a Markov decision process (MDP) using the state spaces, action spaces, and reward functions defined in Phase 1. Across the considered PM decisions, this results in multiple MDP instances with decision-specific state representations, action sets, and reward formulations. Each MDP is subsequently solved by one or more AI models, spanning two extremes, namely separate models per MDP or a single multi-task model across all MDPs. A design of experiments approach is used to identify an efficient configuration, using decision performance based on the Phase 1 KPIs as well as robustness and computational efficiency as evaluation criteria. Reward hacking is prevented by enforcing hard operational constraints at the environment level and by restricting action execution through feasibility checks and action masking, thereby limiting the agent to physically plausible decisions. Policy validation further includes stress tests under randomized disturbances and held-out configurations to detect and penalize strategies that exploit simulation artefacts rather than improving decision-relevant KPIs. However, due to the dynamic evolution of AI research, the final selection of AI-Algorithms must remain flexible and aligned with the state of the art. Alternative approaches, including planning-based methods, generative models, or hybrid architectures, may complement or outperform classical RL techniques. The conceptual structure must therefore allow the integration and comparison of diverse AI methods under identical conditions. Phase 3 thus establishes the foundation for an AI decision system that can navigate and act within the simulation under the same conditions as human players.

Phase 4: Performing the study with human players

Phase 4 focuses on the experimental evaluation of human performance within the simulation. After the simulation environment has been created in phase 2 and the artificial decision system designed in phase 3, real decision-makers are introduced into the system to generate empirical human performance data. The objective is to establish a comprehensive dataset for comparing human and artificial decision-making abilities. A sufficiently large number of participants is required to enable statistically valid conclusions. Based on a t-test with $\alpha = 0.05$ and a medium effect size, approximately 64 datasets per decision type are recommended [46]. While the t-test-based planning suggests approximately 64 datasets per decision type for detecting group-level differences, Phase 5 requires statistically stable item calibration and therefore adopts a Rasch (1PL) IRT model and a precision-oriented, simulation-based IRT sample-size justification, which is expected to require 200 or more samples. Data collection follows principles of empirical social research. Results are anonymized, but optional background information, such as experience or the organizational role can be collected to support interpretation. All participants work through the same simulated decision scenarios as the AI system, enabling direct comparisons. The simulation provides identical information states

and decision options ensuring genuinely comparable testing conditions. Therefore, the results of phase 4 provide the empirical basis on which the evaluation model in phase 5 is built.

Phase 5: Design and application of the evaluation model

The evaluation model in phase 5 integrates the human and artificial performance data collected in the previous phases into a unified analytical framework. This enables a systematic and fair assessment of the reliability of AI-Algorithms in PM decision-making. At the same time, it introduces a methodological challenge, as the simulated decisions can vary considerably in their level of difficulty. Some decisions are simple and well structured, while others are multi-step, uncertain, or highly contextual. Comparing raw scores alone would ignore these differences and result in biased conclusions. A reliable evaluation model must therefore consider both the capability of the decision-makers and the difficulty of the respective decisions. For this reason, phase 5 employs a model-based approach from IRT [47]. IRT enables an simultaneous estimation of latent abilities and decision difficulties, embedding them into a shared reference model. This allows performance differences to be interpreted not only through observed scores but through structural properties of the decision situation. Whether an AI system performs better in simple, deterministic decisions, while humans outperform in complex or unclear scenarios, can be revealed by this technology. In detail, IRT is adapted to the dynamic game environment by converting continuous gameplay into a sequence of anchored decision items with standardized information sets, which supports the local-independence assumption at the measurement level. Context effects arising from evolving system states are addressed through scenario anchoring and by modelling item difficulty as a function of logged state features, so that ability estimates capture decision competence rather than trajectory-specific artefacts. Phase 5 thus provides a scientifically grounded basis for assessing the transferability of specific PM decisions to AI systems.

Phase 6: Real world validation of results

In the final phase, the transferability of the simulation results to real production environments is examined. Although phases 1 through 5 establish a controlled and abstract experimental setting, it is essential to validate the model in practical contexts to assess its real-world reliability. The first step involves discussing the simulated findings with industry experts. The discussion in focus groups allow results to be contextualized, evaluated, and validated against practical experience. Experts help identify discrepancies between model and reality and assess which decision logics are realistic, critical, or highly relevant in practice. In a second validation step, the simulation outcomes are compared with historical production logs from real-world systems to provide an empirical reference and reduce potential circularity arising from expert-based design and validation. The comparison focuses on whether the main performance patterns and KPI trajectories observed in the simulation are also reflected in operational data. A third validation step consists of transferring selected decision problems, especially those where the model predicts a strong advantage for AI or where large differences between human and AI performance were observed, into real production environments. Results obtained in these settings are compared to simulation outcomes to determine whether the observed patterns are reproducible and realistic. Through this two-stage validation, it becomes possible to identify which aspects of the simulation sufficiently reflect real-world decision environments and where refinements are necessary. Phase 6 concludes the benchmark concept and ensures that the developed approach provides a reliable foundation for future assessments of AI's role in PM. In summary, the structured alignment between objectives and phases is summarized in Table 1, outlining how each phase contributes to overall benchmark development. Each phase builds upon insights generated in earlier phases, creating the foundation for a reliable human–AI performance comparison within strategic and tactical decisions in PM.

Table 1: Overview of objectives and corresponding phases

Objective	Phase	Description
A	1	Selecting relevant decisions and developing the simulation environment
	2	

B	3	AI implementation and the execution of the empirical study
	4	
C	5	Human and AI performance evaluation and validation
	6	

4. Initial findings

The initial findings of the underlying research project "Manager ex Machina – Investigation of the transferability of production management decisions to AI systems (ManExMa)" provide a starting point for shaping a human-AI benchmark approach specifically in PM. To clarify which types of PM decisions should be represented in the benchmark, phase 1 has already been initiated. Specifically, 14 expert interviews with production managers from different mechanical engineering companies were conducted and subsequently analysed and synthesized [48]. The consolidated findings were then further refined in consultations with AI experts to incorporate feasibility considerations and expected implementation effort. The resulting task domains with the highest perceived potential for AI adoption are *production controlling*, *process design, financing and investment, operational production management*, and *order management and fulfilment* [48]. These domains form the basis for the subsequent development of the simulation environment. The results of this preliminary study indicate that decisions and challenges within PM are highly similar across companies. At the same time, a shared perspective is emerging regarding which types of decisions are generally suitable for an AI application. However, final and comprehensive conclusions require the conduct and evaluation of the expert interviews. In parallel with these interviews, initial serious-game concepts were analyzed to determine how real decisions from PM can be transformed into a playable and scientifically assessable structure. These preliminary investigations provide the foundation for the conceptualization of an initial benchmark that outlines suitable PM decisions for modelling a serious game. The approach also establishes a game as a scientific method for comparing human and AI performance. The insights gained confirm the basic feasibility of the chosen methodological approach and form the basis for the next development steps.

5. Conclusion, limitations and further research

The development of the simulation-based benchmark concept provides a novel approach for systematically assessing the decision-making capabilities of AI in PM. By transferring benchmarking principles from established game environments into an abstracted and realistic production context, the project introduces a conceptual foundation that enables a controlled, reproducible, and valid performance comparison between human and artificial decision-makers. This paper answers the underlying research question of how a standardized, controllable, and repeatable methodology can be developed to objectively compare human and AI decisions in PM. This approach is relevant across domains, as many industrial and organizational settings face similar decision-making challenges characterized by complexity and uncertainty. The resulting benchmark will support companies in evaluating which types of decisions can reliably be supported or potentially taken over by AI systems in the future, and in which areas human expertise will remain indispensable. At the same time, it will reduce the need for time-consuming, cost intensive, and risky experimentation within real production operations. The objective is not to replace human decision-makers, but to identify synergistic potential that can make production systems more efficient, robust, and adaptable. The simulation-based approach provides a risk-free environment, enabling the testing of AI systems without impacting ongoing processes or operational performance indicators. As the underlying research project ManExMa is still in an early stage of development, its methodological foundations will be continuously refined. The next steps involve further detailing the simulation and the evaluation model and enhancing both with the results from expert interviews and human gameplay sessions. Moreover, the project will examine how differences in decision complexity and difficulty levels can be represented consistently and in a scientifically valid manner within the evaluation model. A particular focus of future work should be the

validation of the approach in real production environments. This includes expert assessments, the validation against historic production logs and the exemplary implementation of selected decision scenarios in real production processes. In addition, the organizational, economic, and ethical implications of an increased use of AI in PM will be examined. Overall, the conceptual benchmark developed in this project may lay the foundation for a standardized and transferable evaluation methodology that, for the first time, could enable a systematic and comparable assessment of AI performance in complex decision-making domains.

Acknowledgements

This paper, as well as others, are published within the research project "ManExMa" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 525213887, which results in overlapping documentation and research based on previous publications. Hence similarities between this paper and preceding papers published by the chair of IPEM are to be expected.

References

[1] Lachova, K., 2021. Digital Transformation as a Strategy of Value Creation in Industry. Acta Mechanica Slov. 25 (3), 52–56.

[2] Broy, M., 2010. Cyber-Physical Systems. Springer Berlin Heidelberg, Berlin, Heidelberg.

[3] Lee, J., Bagheri, B., Kao, H.-A., 2015. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters 3, 18–23.

[4] Spath, D., Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T., Schlund, S., 2013. Produktionsarbeit der Zukunft - Industrie 4.0.

[5] Raisch, S., Krakowski, S., 2021. Artificial Intelligence and Management: The Automation–Augmentation Paradox. AMR 46 (1), 192–210.

[6] Wuest, T., Weimer, D., Irgens, C., Thoben, K.-D., 2016. Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research 4 (1), 23–45.

[7] Gröger, C., 2015. Advanced Manufacturing Analytics. Josef Eul Verlag, Lohmar.

[8] Pokorni, B., Braun, M. & Knecht, C., 2021. Menschenzentrierte KI-Anwendungen in der Produktion: Praxiserfahrungen und Leitfaden zu betrieblichen Einführungsstrategien. Fraunhofer IAO.

[9] Mockenhaupt, A., 2021. Digitalisierung und Künstliche Intelligenz in der Produktion. Springer Fachmedien Wiesbaden, Wiesbaden.

[10] Burggraf, P., Wagner, J., Koke, B., 2018. Artificial intelligence in production management: A review of the current state of affairs and research trends in academia, in: 2018 International Conference on Information Management and Processing (ICIMP). 2018 International Conference on Information Management and Processing (ICIMP), London. 12.01.2018 - 14.01.2018. IEEE, pp. 82–88.

[11] Sauer, C.R., Burggräf, P., Steinberg, F., 2025. A systematic review of machine learning for hybrid intelligence in production management. Decision Analytics Journal 15, 100574.

[12] Dyckhoff, H., Spengler, T., 2007. Produktionswirtschaft. Springer Berlin Heidelberg, Berlin, Heidelberg.

[13] Nassehi, A., 2018. Operations Management, in: , CIRP Encyclopedia of Production Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–7.

[14] Omar, Z.S., Bo, H., 2022. A Company Production Management Optimization Research. AJIBM 12 (03), 365–375.

[15] Schönsleben, P., 2016. Integrales Logistikmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg.

[16] Bashiri, M., Badri, H., Talebi, J., 2012. A new approach to tactical and strategic planning in production–distribution networks. Applied Mathematical Modelling 36 (4), 1703–1717.

[17] Sauer, C.R., 2026. Human-AI Collaboration in Production Management. Springer Fachmedien Wiesbaden, Wiesbaden.

[18] Urgo, M., Lanza, G., Vrabič, R., Gyulai, D., 2025. Future-Proof Production Scheduling and Control. CIRP Annals 74 (2), 971–991.

[19] Carvalho, A.N., Scavarda, L.F., Oliveira, F., 2017. An optimisation approach for capacity planning: modelling insights and empirical findings from a tactical perspective. Prod. 27 (0).

[20] Seshia, S.A., Sadigh, D., Sastry, S.S., 2022. Toward verified artificial intelligence. Commun. ACM 65 (7), 46–55.

[21] Zhang, L., Pan, Y., Wu, X., Skibniewski, M.J. (Eds.), 2021. Artificial Intelligence in Construction Engineering and Management. Springer Singapore, Singapore.

[22] Russell, S.J., Norvig, P., 2010. Artificial intelligence: A modern approach, 3. ed., international ed. ed. Pearson, Boston, Munich, 1132 pp.

[23] Raschka, S., Mirjalili, V., 2021. Machine Learning mit Python und Keras, TensorFlow 2 und Scikit-learn: Das umfassende Praxis-Handbuch für Data Science, Deep Learning und Predictive Analytics, 3., aktualisierte und erweiterte Auflage ed. mitp, Frechen, 770 pp.

[24] Cummings, M.M., 2014. Man versus Machine or Man + Machine? IEEE Intell. Syst. 29 (5), 62–69.

[25] Fitts, P.M., 1951. Human engineering for an effective air-navigation and traffic-control system. National Research Council Division of Anthropology and Psychology Committee on Aviation Psychology, Washington, 84 pp.

[26] Sheridan, T.B. Function allocation: algorithm, alchemy or apostasy? International Journal of Human-Computer Studies 2000 (52), 203–216.

[27] Henriksen, D., Creely, E., Gruber, N., Leahy, S., 2025. Social-Emotional Learning and Generative AI: A Critical Literature Review and Framework for Teacher Education. Journal of Teacher Education 76 (3), 312–328.

[28] Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. The M4 Competition: Results, findings, conclusion and way forward. International Journal of Forecasting 34 (4), 802–808.

[29] Reichenstein, T., Raffin, T., Sand, C., Franke, J., 2022. Implementation of Machine Vision based Quality Inspection in Production: An Approach for the Accelerated Execution of Case Studies. Procedia CIRP 112, 596–601.

[30] Ni, S., Chen, G., Li, S., Chen, X., Li, S., Wang, B., Wang, Q., Wang, X., Zhang, Y., Fan, L., Li, C., Xu, R., Le Sun, Yang, M., 2025. A Survey on Large Language Model Benchmarks.

[31] Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., Hassabis, D., 2017. Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm.

[32] Vinyals, O., Babuschkin, I., Czarnecki, W.M. et al., 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature 575 (7782), 350–354.

[33] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518 (7540), 529–533.

[34] Brynjolfsson, E., Mitchell, T.M., 2017. What can machine learning do? Workforce implications. Science, 1530–1534.

[35] Gąsior, D., Józefczyk, J., 2009. Application of uncertain variables to production planning in a class of manufacturing systems. Bulletin of the Polish Academy of Sciences: Technical Sciences 57 (3).

[36] Yin, K., Liu, H., Yin, G., 2003. Stochastic models and numerical solutions for production planning with applications to the paper industry. Computers & Chemical Engineering 27 (11), 1693–1706.

[37] Dulac-Arnold, G., Mankowitz, D., Hester, T., 2019. Challenges of Real-World Reinforcement Learning.

[38] Rusdi, H., Kusumasari, I.R., Putri, P.A., Murdiana, N., Puspita, D.R., 2024. Challenges and Opportunities for Using Artificial Intelligence as a Supporting Tool in Business Decision Making in the Digital Era. jbkd 2 (2), 17.

[39] Wang, L., Jiang, Z., Hu, C., Zhao, J., Zhu, Z., Chen, X., Wang, Z., Liu, T., He, G., Yin, Y., Lee, D.-H., 2025. Comparing AI and human decision-making mechanisms in daily collaborative experiments. iScience 28 (6), 112711.

[40] Ovezmyradov, B., 2025. AI Playing Business Games: Benchmarking Large Language Models on Managerial Decision-Making in Dynamic Simulations.

[41] Lalor, J.P., Wu, H., Yu, H., 2016. Building an Evaluation Scale using Item Response Theory. Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing 2016, 648–657.

[42] Wu, K., 2014. A Unified View on Planning, Scheduling and Dispatching in Production Systems.

[43] Aghamanoukjan, A. Quantitative Interviews, in: Renate Buber und Hartmut H. Holzmüller (Ed.), Qualitative Marktforschung. Konzepte - Methoden - Analysen, 2 ed. Verlag Gabler, Wiesbaden, pp. 415–436.

[44] Mayring, P., Fenzl, T., 2014. Qualitative Inhaltsanalyse, in: , Handbuch Methoden der empirischen Sozialforschung. Springer VS, 2014, Wiesbaden.

[45] Sutton, R.S., Barto, A., 2018. Reinforcement learning: An introduction, Second edition ed. The MIT Press, Cambridge, Massachusetts, London, England, 590 pp.

[46] Döring, N., Bortz, J., 2016. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften, 5. vollst. überarb. Aufl. ed. Springer, Berlin, 1051 pp.

[47] Wirth, R.J., Edwards, M.C., 2007. Item factor analysis: current approaches and future directions. Psychological methods 12 (1), 58–79.

[48] Stracke, L., Burggräf, P., Sauer, C.R., Schütz, M., Kaiser, J., 2026. From human to machine: high-impact tasks for AI in production management – an expert study to reshape decision-making. Prod. Eng. Res. Devel. 20 (1).

Biography

Joshua Kaiser, born 1997, studied Mechanical Engineering and International Production Engineering and Management (IPEM) at University of Siegen. Since 2022, he has been a Research Associate with the IPEM, University of Siegen and is pursuing his Ph.D. degree. His research focuses on the role of artificial intelligence in production management, particularly in simulation-based performance evaluation.

Carl René Sauer was born in 1997. He studied Industrial Engineering and International Production Engineering and Management at University of Siegen. Since 2019, he has been a Research Associate with the IPEM, University of Siegen, and since 2025 he is chief engineer at the chair for International Production Engineering and Management (IPEM). He finished his Ph.D. degree in 2025, while the focus of his research projects includes hybrid intelligence in production management, circular economy and ML.

Norman Müller was born in 1996. He studied Industrial Engineering at the University of Siegen and completed his M.Sc. degree in 2022. Since 2022, he has been a research associate at IPEM, University of Siegen, where he is pursuing his Ph.D. and since 2025, he is group lead for Intelligent Production and Supply Chain Management. His Ph.D. focuses on the prediction of delivery delays for purchased parts in manufacturing companies using artificial intelligence.

Luisa Stracke, born in 1999, studied Industrial Engineering at the University of Siegen, where she completed both her bachelor's and master's degrees. She is currently pursuing her Ph.D. at the Chair of International Production Engineering and Management (IPEM) at the University of Siegen. Her research focuses on the systematic investigation of the extent to which artificial intelligence can take over production management tasks in industrial environments.

Peter Burggräf, born 1980, studied Mechanical Engineering in Aachen and London. He wrote his PhD thesis at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University in the field of factory planning. Since 2017, Prof. Burggräf holds the chair for International Production Engineering and Management (IPEM) at the University of Siegen in Germany.

8th Conference on Production Systems and Logistics

Structuring Design Problems of Reconfigurable Robotic Assembly Systems

Elias Huber, Marc Ungen
Corporate Research/ Robert Bosch GmbH, Renningen, Germany

Abstract

A Reconfigurable Robotic Assembly System (RRAS) is an assembly system employing industrial robots configured as reconfigurable modular machines. Determining the optimal assembly of a particular product within an RRAS presents a complex design problem. This problem is characterized by two primary dimensions: (1) the determination of the system's physical structure and (2) the definition of the assembly procedures. The combined structural and procedural configuration space is extensive and exhibits interdependencies. Existing research approaches address only isolated subproblems and fail to provide a comprehensive perspective on the overall design challenge. The contribution of this paper is to fill this gap by systematically structuring the RRAS design problem. Through a systematic literature review, we identify aspects that influence the design in the selected publications. We subject a study case to hierarchical decomposition to break down the design into distinct problems. The interdependencies among these individual problems are then mapped into a hierarchical structure to provide a comprehensive representation. The resulting structure assists researchers in defining and clarifying the subproblem they are addressing. This improves comparability across studies and serves as template for guiding integrated design and optimization of RRASs.

Keywords

Robotics; Optimization; Reconfigurable System; Assembly System Design; Scheduling; Factory Automation

1. Introduction

Reconfigurable manufacturing systems are regarded as a promising solution to meet rapid market changes [1, 2]. According to the definition provided by Koren et al. [3], reconfigurable modular machines form the hardware component of reconfigurable manufacturing systems. This study focuses on assembly systems and on industrial robots. These robots can change their tools and can therefore adapt their capabilities. They can act as reconfigurable modular machines. To denote the specific type of reconfigurable manufacturing system considered in this work, we introduce the term Reconfigurable Robotic Assembly System. Koren et al. [3] identify the time intensive design of such systems as a major obstacle. This challenge is critical because reconfigurable systems must be redesigned often to meet changing requirements. To automate the design, researchers formulate the selection among different design options as computational problems. Examples include the number of robots, the positioning of robots, and the assignment of tasks. These problems are solved through computation to obtain feasible or optimal solutions [4, 11, 15, 16]. We group the individual design problems along two dimensions. These are the physical structure of the system and the procedure of

DOI: https://doi.org/10.15488/20987
ISSN: 2701-6277

the assembly processes. We define the terms physical structure, procedure and process in accordance with the IEC standard 61512 (ISA-88) standard [5].

1. A physical structure includes the hardware of a system such as robots and tools. It defines the physical capabilities of the system such as workloads or motion accuracy.
2. A process is an activity that changes the geometry of a product. Screwing, dispensing and inserting are typical examples. Processes can include constraints to ensure quality such as required trajectories or torque values.
3. A procedure defines the strategy for carrying out a process. It specifies which robot executes the task, which tool is used, and which motions are required.

The objective of an assembly task is to find a procedure that satisfies the process constraints and matches the physical capabilities of the system. A configuration describes the structure of the system, and the procedures used to assemble a product. Determining the optimal configuration for a reconfigurable system is difficult [3]. The term optimal configuration is broad and may refer to criteria such as throughput, cost or energy use. These criteria may conflict with each other. The number of possible configurations grows rapidly with the number of robots. Evaluating each configuration is not practical because this requires extensive computational effort. The configuration space is also difficult to navigate. Changes in the structure of the system or in the assembly procedures can strongly influence the best configuration in the other dimension. Recent research has addressed many specific subproblems of the overall design task [4, 9, 10, 14, 15]. However, these works do not provide a complete view of the full design problem. The objective of this paper is therefore to break the overall design problem into individual problems and to describe their interdependencies. We structure these problems hierarchically. The resulting structure serves as a template. It helps researchers identify which problems they address, and which problems lie outside the scope of their work. This clarity supports the identification of further optimization potential or design alternatives. The structure also provides a shared reference for comparing different approaches and for highlighting similarities and differences between them.

1.1 Methodical Approach

Goel et al. [6] describe a four-step process for solving a general design problem. The first two steps analyse the problem. The last two steps solve it. This work focuses only on the analysis phase, since our goal is to build a clear understanding of the overall problem. The two analysis steps are:

1. An exploration and decomposition of the problem.
2. An identification of the interconnections among the components.

Because we do not study a specific design case, we first derive a representative problem that reflects the design of RRAS. To do this, we conduct a systematic literature review in Section 2.1. We query the database Scopus with a defined search string and exclude irrelevant publications. We then analyse the selected works to extract the structural and procedural aspects that shape the RRAS design problem. Using these aspects, we define the representative problem in Section 2.2. In Section 3.1, we decompose this problem into individual problems across three abstraction levels. At this stage, each problem is treated independently. In Section 3.2, we perform the second analysis step and describe how the individual problems are linked. Section 4 shows how the resulting structure can serve as a template for specifying the design problems studied in existing publications.

2. Exploration of the Design Problem

2.1 Literature Review

The search string *Reconfigurable AND Robotic AND Assembly AND (System OR Cell)* was applied in Scopus for journal articles and conference papers. The final query was executed on December 10, 2025, with the

search restricted to publications from 2010 onward. Publications were screened using a multi-stage process comprising title review, abstract review, an initial full-text scan, and in-depth reading. At each stage, papers were excluded if they described the same system multiple times, lacked sufficient architectural detail, or focused on domains outside the scope of this study, such as communication or control systems, mobile robotics, specific manufacturing processes, aerospace, construction, robot design, or human-robot collaboration. Eleven publications remained for analysis. These works contain eleven recurring aspects. Seven aspects influence the structural configuration of an RRAS:

1. Multiple Stages MS - Consists of multiple distinct units that are spatially separated and execute a delimited set from the overall tasks.
2. Industrial Robot IR - At least one industrial robot is used and simulated.
3. Multiple Industrial Robots MIR - More than one industrial robot is used and simulated.
4. Mobile Robot MR – At least one mobile robot is used and simulated.
5. Layout Configuration LC - The positioning of the hardware components is adaptable.
6. Crowded Workspace CW - The industrial robots are positioned so that their workspaces overlap, creating a highly dynamic collision scene.
7. Resource Selection RS - The hardware components of the system are selected and not predefined.

Four aspects influence the procedural configuration of an RRAS:

1. Resource Configuration RC - The capabilities of robots or machines are adapted to adjust them to different tasks.
2. Task Assignment TA - Tasks are assigned based on their requirements.
3. Action Generation AG - Tasks are sequenced in discrete logical steps, which we refer to as actions.
4. Parallelization of Tasks or Actions PTA – Tasks or actions execution is parallelized on different machines or robots.

Multiple stages, industrial robots and multiple industrial robots define the fundamental structural outlines of the systems. Across the eleven publications, five combinations of these aspects are identified. A combination of multiple stages and multiple robots is described in [7, 8, 9, 10]. Up to two stages with one industrial robot each are investigated in [14]. Systems with a single stage but with multiple industrial robots are described in [11, 12]. A configuration with a single cell and a single industrial robot is found in [13, 14]. Systems with multiple stages but without specific robot simulation, instead defining black box resources with assigned capabilities, are presented in [15, 16, 17]. Mobile robots and layout configuration appear in different forms. In [14] mobile robots are considered but not fully integrated into the system design, whereas [7] describes mobile articulated robots that reposition themselves to handle tasks. A more static layout appears in [8, 9, 10, 12, 14, 16], where components are positioned once before the assembly procedure begins. The crowded workspace aspect requires that several robots operate within the same stage. Publications [9, 10] incorporate multiple stages but restrict each stage to a single robot, and [12] uses manual repositioning and teach in motions, which leads to a neutral evaluation of layout configuration. As a result, only [7, 11] implement several industrial robots within one stage and therefore include the crowded workspace aspect. Resource selection is addressed in almost all publications; only [12, 13, 17] rely on predefined component sets and therefore apply a bottom-up approach. Resource configuration is treated in all eleven publications, since adapting robot or machine capabilities is essential in reconfigurable systems. Task assignment is addressed in [7, 9, 10, 11, 12, 13, 16], while the remaining works use predefined sequences. Action generation is considered in [9, 10, 11, 13, 14, 16], where robots derive the required actions from their current state or from necessary tool changes. Parallelized tasks or actions are examined in systems using black box machines in [16, 17], multiple robots in [8, 11] and integrated robot machine systems in [13]. The publication by Michalos et al. [7], titled “The New Assembly System Paradigm”, provides a high-level overview of a prospective manufacturing system without an in-depth analysis of the associated design

challenges. Although aspects such as action generation and parallelization would likely be relevant, they are not explicitly discussed. Table I presents the comprehensive overview.

Table 1: Design aspects of RRAS in publications

Publication	MSS	MIR	IR	MR	LC	CW	RS	RC	TA	AG	PTA
E. Carpanzano et al. (2011) [13]	✗	✗	✓	✗	✗	✗	✗	✓	✓	✓	✓
J. Padayachee et al. (2012) [16]	✓	✗	✗	✗	✓	✗	✓	✓	✓	✓	✓
G. Michalos et al. (2015) [7]	✓	✓	✓	✓	✓	✓	✓	✓	✓	–	–
F. G. Sisca et al. (2015) [17]	✗	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗
C. Schlette et al. (2020) [12]	✗	✓	✓	✗	–	–	✗	✓	✓	✗	✗
F. Shi et al. (2021) [8]	✓	✓	✓	✗	✓	✗	✗	✓	✗	✗	✓
C. Zhang et al. (2023) [10]	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✗
M. Albus et al. (2024) [15]	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗
M. Ungen et al. (2024) [11]	✗	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓
M. Uzunosmanoglu et al. (2024) [14]	✓	✓	✓	–	✓	✗	✓	✓	✗	✓	✗
Z. Xu (2024) et al. [9]	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✗

2.2 Derivation of a Representative Design Problem

By defining a design problem that integrates all relevant aspects we assume that all underlying individual problems are included. RRAS design problems that omit certain aspects can be regarded as subclasses that still follow the structure developed in this paper. The configuration in [7] covers many aspects, but it places a strong emphasis on mobile articulated robots and is therefore distinct from the other studies. To define a representative design problem for all analysed publications we propose a configuration that balances the dynamic layout in [7] with the more static layouts found in the remaining works. Therefore, the representative system under investigation consists of two parallel rows of articulated robots arranged in a staggered layout with overlapping workspaces, as illustrated in Figure 1. Each robot is equipped with a tool set that enables the execution of different assembly processes. The robots are grouped into stages that contain one robot, two robots or three robots. These stages form spatially separated units. The workspace of each robot is restricted to the borders of its stage so that collision avoidance is handled locally. Material flow between stages is ensured by mobile workpiece carriers that also establish movable process stations and deliver components where needed. The next abstraction level is the module level which represents a single robot. The system level, the stage level and the module level form the hierarchical structure used for the decomposition of the design problem as shown in Figure 2. The system is designed to perform three types of assembly tasks which are screwing, insertion and dispensing. These tasks reflect common activities in assembly technology. Process parameters such as trajectories and torque limits are predefined to ensure quality. Handling motions must be generated for the specific scene. The resulting design problem is to find a configuration that maximizes throughput while remaining within a given cost limit.

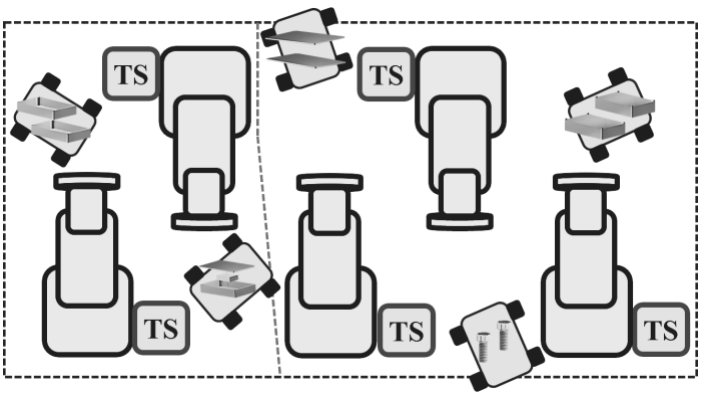


Figure 1: Exemplary instance of the representative RRAS

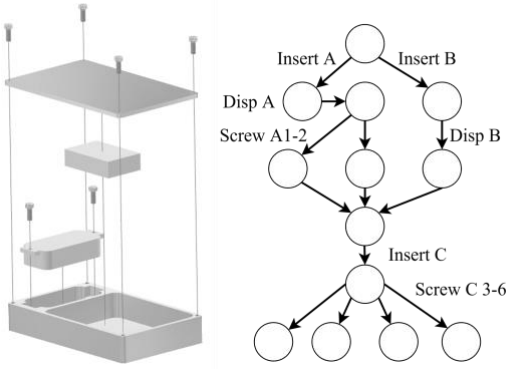


Figure 2: Representative set of assembly tasks depicted as precedence graph

3. Structuring of the Design Problem

3.1 Decomposition into Individual Problems

In this section we identify the individual problems at the system level, the stage level and the module level by applying a hierarchical decomposition to the representative problem. The approach begins with high level considerations and becomes more detailed with each level. The objective is to achieve a finer resolution of the overall problem structure step by step.

At the system level three main design considerations define the individual problems. These are the definition of the general structure of the system, the segmentation of the problem into manageable subparts and the alignment of these subparts. These considerations create the abstract outline of the system and define the boundaries for the lower abstraction levels. The general structure results from the selection of the hardware and their spatial arrangement. The representative design problem already defines some of these aspects. The hardware includes articulated robots with tool sets and mobile robots. The problem of identifying the type and quantity of hardware without specifying the exact robot models is called *System Resource Selection*. The basic arrangement of the hardware is defined by *System Layout Definition*. The second main consideration assumes that complex problems must be divided into smaller parts. The *Definition of Stages* describes how the system is segmented. The distribution of the assembly tasks to the stages is called *Task Assignment to Stages*. The alignment of the stages must also be considered. To maximize throughput all stages should require similar processing times. This is addressed by *Stage Balancing*. The individual system level problems are shown in the left column of Figure 3.

To proceed from the system level to the stage level we define a representative stage. The representative stage, shown in Figure 4, contains three articulated robots that execute a subset of the assembly tasks. A stage represents a smaller subproblem and therefore allows a more detailed specification. Some system level considerations appear again but must be refined. The first consideration defines the exact structure of the stage. The second consideration concerns the alignment of the modules within the stage. These two aspects define the stage level problems. *Resource Selection* must now identify the exact robot models. *Layout Definition* must define the absolute position of the robots and their tool sets. Mobile robots act as movable process stations and workpiece carriers. Their positions and pick up locations are also defined in the layout. *Task Assignment* to modules includes two steps. First, the assembly tasks are assigned to the articulated robots. Second, the transport tasks between process stations and pick up locations are assigned to the mobile robots. Aligning modules into a cohesive stage is different from aligning stages within a system. Robots in one stage have overlapping workspaces. Collision avoidance must therefore be satisfied. The motions of the mobile robots must also be coordinated. They must match the process times of the articulated robots. These requirements are addressed by *Stage Motion Planning*. The temporal and logical synchronization of the modules is the goal of the *Scheduling* problem. This problem optimizes the execution times while respecting the assembly precedence graph. The stage level problems are shown in the central part of Figure 3.

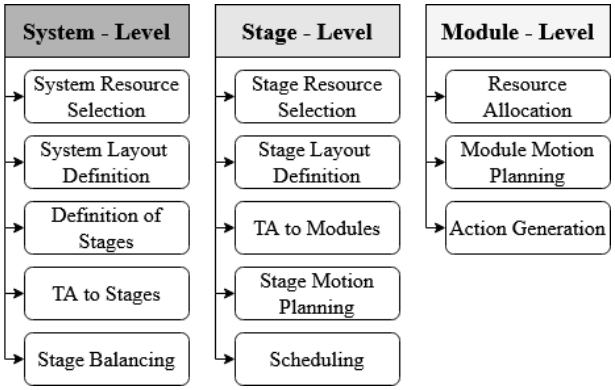


Figure 3: Individual problems at the three levels of abstraction

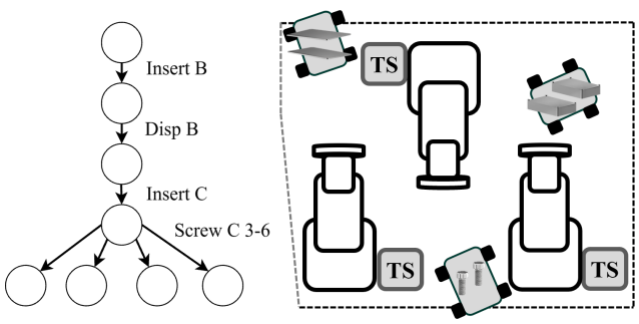


Figure 4: Subset of the assembly tasks and exemplary stage

At the module level the goal is to determine the detailed procedure for each task assigned to a robot. Three considerations describe this level. These are the allocation of tools to robots and tasks, called *Resource Allocation*. The definition of the logical steps required for task execution and the motion of the robot during execution. Matching tasks with the correct tool is addressed by resource allocation. *Action Generation* creates the sequence of steps needed for task execution. For example, a screwing task may include tool selection, screw retrieval, positioning and the screwing action. *Module Motion Planning* calculates the motions required for each step. The module level problems are shown in the right column of Figure 3.

3.2 Interconnections of the Individual Problems

We examine the interconnections among the individual problems while preserving the distinction between the three levels of abstraction. Since a stage includes a subproblem, multiple stage instances exist. For simplicity, Figure 5 depicts one stage in detail and includes another schematic stage as a placeholder for additional instances. The same approach applies to modules that collectively form a stage. The dependencies between the individual problems are represented by unidirectional or bidirectional arrows.

System Resource Selection, System Layout Definition, Definition of Stages, and *Task Assignment to Stages* shape the fundamental system design. Decisions about resource selection and the basic layout configuration mutually influence each other. Likewise, the selection and arrangement of resources directly can impact the optimal segmentation of resources and tasks. Additionally, the *Definition of Stages* and *Task Assignment* to Stages can influence which resources are needed and how they should be spatially arranged. For example, determining that the system must have three stages might require a completely different resource selection and arrangement than a system with four stages. Therefore, these four problems and their respective decisions mutually influence each other. The Task Assignment to Stages problem defines the procedural system design, while the other three influence the system's structural design. The effects on the design dimensions are depicted in Figure 5 by unidirectional arrows leading to structural or procedural design storage nodes. Input for the *Stage Balancing* problem are the cycle times of the stages. Within this problem, it is evaluated whether modifications to the system configuration can improve the system's throughput. To initiate these modifications, further constraints to the individual problems can be made to alter their solution space. The adaptation of constraints can be made directly to system-level problems or to stage-level problems. The dependencies that arise by the influence of the Stage Balancing are indicated by dashed arrows.

The stage-level problems *Stage Resource Selection*, *Stage Layout Definition*, and *Task Assignment* to Modules are strongly influenced by the structural and procedural design decisions established at the system level. Furthermore, they are mutually dependent as the following example points out. The absolute positioning of two robots determines their distance. If both are assigned tasks that must be executed at the same process station, the workspaces of the specific robot models used must overlap sufficiently to ensure they can reach and perform the required motions. The execution time of a task depends on whether a collision-free motion can be planned at that moment. Thus, *Scheduling* depends on *Stage Motion Planning* to ensure feasibility. Conversely, the set of possible motions at any given time is constrained by the tasks scheduled for the other robots in the stage. Therefore, *Stage Motion Planning* and *Scheduling* are mutually dependent. If the outcome of the *Scheduling* problem indicates that the modules cannot be aligned in a way that meets the system's requirements, additional constraints are imposed on the stage- or module-level problems. These dependencies that arise from the necessity to align modules within a stage are depicted by dashed arrows in Figure 5.

The individual problems at module-level *Resource Allocation*, *Module Motion Planning* and *Action Generation* are constrained by stage-level design decisions. Furthermore, *Resource Allocation* influences both *Module Motion Planning* and *Action Generation* as the following examples show. A different robot-tool matching can require different motions because of different geometries. The application of a multi-spindle screwdriver in comparison to a single-spindle driver can reduce the number of actions needed.

Therefore, the *Resource Allocation* can influence the other two problems. In addition, if motions or actions are predefined the tool selected must conform to these constraints. Lastly, the definition of the action sequence influences the sequence of motion, and the actions needed depend on whether feasible motions can be calculated. Therefore, the three module-level problems are mutually dependent. The resulting module procedures are input to the Scheduling problem.

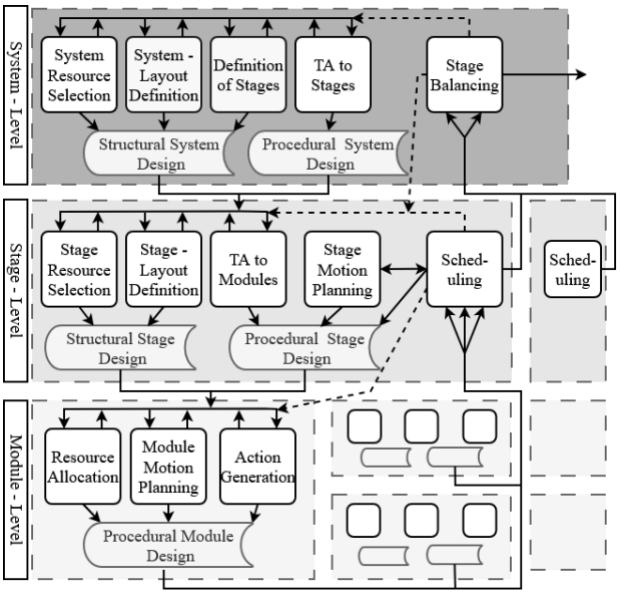


Figure 5: Hierarchical structure of the design problem

4. Application

The hierarchical structure is established to address all individual problems that must be considered to achieve an optimal system configuration. However, in both research and engineering practices, system designers may not always be able to consider each individual problem in detail. Moreover, certain problems are predetermined by the design prerequisites. In such instances, the problem structure can be adapted to the specific design problem. To exemplify this, we will analyse the publications [11], [14] and adjust the hierarchical structure to the design problem defined and solved there. This shows their similarities and differences in defining the design problem and makes their approaches comparable. [11] defines a scenario involving one to three six-axis robotic manipulators for the assembly of one or two identical products. The objective is to minimize assembly time by identifying a task assignment that enables the robots to operate simultaneously on the same products. In this context, all robots share portions of their workspaces. Consequently, the design problem does not address problems at the system level. It can be directly constrained to a single stage consisting of one to three modules. At the stage-level, the problem is further restricted in its structural dimension. The solution space for Stage Resource Selection is constrained to one to three six-DoF robotic manipulators of the same model. Additionally, the Stage Layout problem is not addressed, as the positions of these robots are fixed and cannot be altered. The adjusted structure according to [11] is depicted in Fig. 6 on the left. In this representation, individual problems that are not addressed by the setup have been omitted, and those whose solutions are predefined by scenario-based definitions are shown in grey. The solution approach employed by [11] also affects the dependencies of the individual problems discussed. In one optimization run, the method preselects the number of robots to be used, optimizes the task assignment, and then compares the solution with other runs that employ different numbers of robots. Consequently, Task Assignment to Modules depends on Resource Selection, but not vice versa. At the module-level, for a task sequence, the actions and the resource allocation are defined in a first step. The Module Motion Planning is executed as a subsequent step. The motions of the individual modules are sequenced to ensure that the robot motions do not collide with each other. Optimization is then achieved by adjusting the task assignment to the modules. In comparison, [14] examine a scenario comprising up to two

stages. In the centre of each stage is a single industrial robot. The robot can be of two possible types. Around the robot, optional add-on modules can be arranged in a hexagonal cell structure. These add-ons include basic stations, transportation units, bin-pickers, and tool warehouses. The system is designed to produce a changing product mix across multiple periods. The planning objective is to minimize costs by jointly optimizing the selection of resources and the routing of products through the system. The key distinction from the scenario described in [11] lies in the incorporation of multiple stages, albeit restricted to one robot per stage. As a result, the task assignment problem is realized at the system-level. Furthermore, the nature of the motion planning problem differs significantly. [14] reduce complexity by excluding the need to coordinate and schedule motions between multiple robots within a stage. This allows precomputed motions to be reused during planning. At the same time, the resource selection problem becomes more advanced, since the optimization must also determine the add-ons to be deployed in each period. In contrast, in [11] these resources are predefined. The structure of the problem according to [14] is depicted in Figure 6 on the right. In summary, both [11] and [14] contribute to the design of reconfigurable robotic assembly systems, yet their distinct scopes lead to different problem definitions. While [11] concentrates on task assignment and multi-robot motion planning within a single stage, [14] extends the scope to multiple stages and emphasizes resource selection and planning across multiple periods. Representing the design problems within the hierarchical structure enables a direct comparison of the aspects each approach addresses.

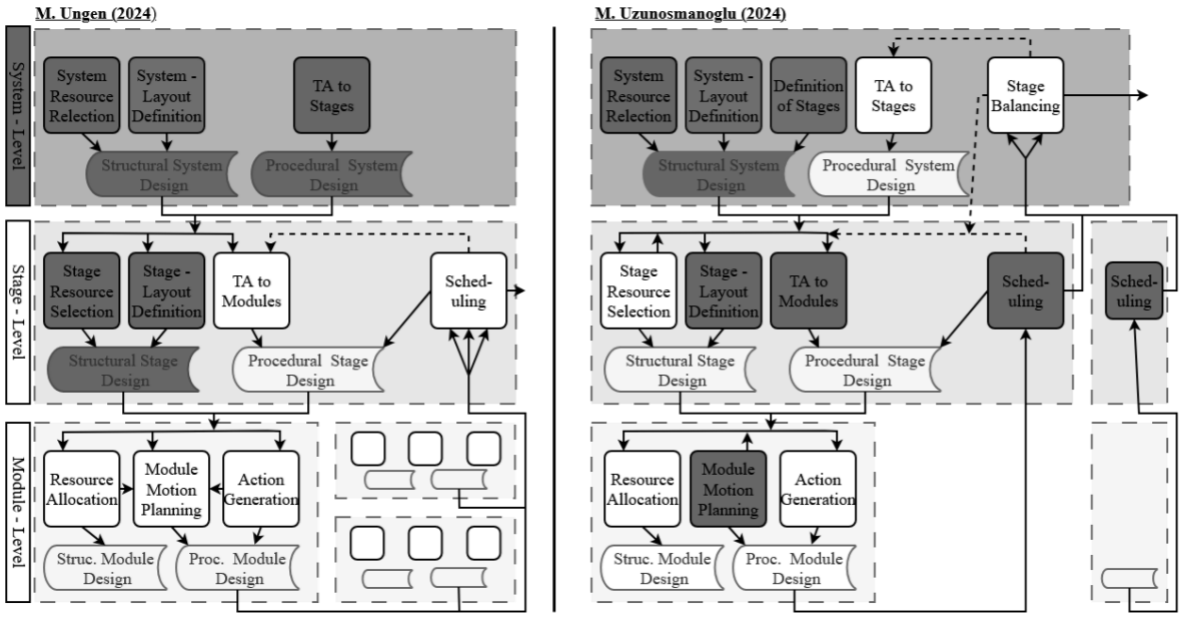


Figure 6: Adaptation of the structure to describe the design problems in [11] on the left and in [14] on the right

5. Conclusion

This paper structured the design problem RRASs. The problem was decomposed into subproblems, and their interconnections were mapped across three abstraction levels: system, stage, and module. The hierarchy provides a perspective that can be applied to clarify the scope of studies and to make assumptions explicit. The contribution is a framework that covers both structural and procedural aspects of RRAS design. It allows researchers to position their work within the overall problem space. Overlaps and differences between approaches become visible. The framework also indicates whether a method is limited to specific assumptions or can be applied more generally. For practitioners, it offers a template to analyse system configurations. It makes clear which aspects of the system are fixed constraints and which remain open for decisions. The structure is conceptual and not yet validated in industrial applications. Future work should test it in real design scenarios. Another step is to link the identified subproblems to formal optimization problems. Existing approaches in the literature address many of these subproblems in isolation.

References

[1] Azab, A., ElMaraghy, H., Nyhuis, P., Pachow-Frauenhofer, J., Schmidt, M., 2013. “Mechanics of change: A framework to reconfigure manufacturing systems.” CIRP Journal of Manufacturing Science and Technology 6 (2), 110–119.

[2] Koren, Y., 2010. The Global Manufacturing Revolution, 1st ed. John Wiley and Sons, New York, USA.

[3] Koren, Y., Heisel, U., Jovane, F., Moriwaki, T., Pritschow, G., Ulsoy, G., Van Brussel, H., 1999. “Reconfigurable manufacturing systems.” CIRP Annals – Manufacturing Technology 48 (2), 527–540. doi: 10.1016/S0007-8506(07)63232-6.

[4] Ungen, M., Kampert, D., Feldotto, B., Huber, E., Riedel, O., 2024. “Automated workflow generation supporting the value stream design of reconfigurable robot assembly cells.” Procedia CIRP 128, 609–614.

[5] International Electrotechnical Commission, 1997. Batch Control – Part 1: Models and Terminology. IEC Standard 61512-1.

[6] Goel, V., Pirolli, P., 1992. “The structure of design problem spaces.” Cognitive Science 16 (3), 395–429.

[7] Michalos, G., Makris, S., Chryssolouris, G., 2015. “The new assembly system paradigm.” International Journal of Computer Integrated Manufacturing 28 (12), 1252–1261.

[8] Shi, F., Wang, S., Zeng, L., 2021. “The balancing problem for a reconfigurable flexible assembly line.” In: Proceedings of the 6th International Conference on Automation Control and Robotics Engineering (CACRE 2021), 392–397.

[9] Xu, Z., Zhang, C., Hu, S., Han, Z., Feng, P., Zeng, L., 2024. “Reconfigurable flexible assembly model and implementation for cross-category products.” Journal of Manufacturing Systems 77, 154–169.

[10] Zhang, C., Dou, J., Wang, P., 2023. “Configuration design of reconfigurable single-product robotic assembly line for capacity scalability.” Computers and Industrial Engineering 185, 109682.

[11] Ungen, M., Huber, E., Kampert, D., Riedel, O., 2024. “Assembly procedure generation for reconfigurable robot cells considering operation concurrency and geometrical constraints.” In: Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE), 981–987. doi: 10.1109/CASE59546.2024.10711657.

[12] Schlette, C., Buch, A.G., Hagelskjær, F., Iturrate, I., Kraft, D., Kramberger, A., Lindvig, A.P., Mathiesen, S., Petersen, H.G., Rasmussen, M.H., Savarimuthu, T.R., Sloth, C., Sørensen, L.C., Thulesen, T.N., 2020. “Towards robot cell matrices for agile production – SDU Robotics’ assembly cell at the WRC 2018.” Advanced Robotics 34 (7–8), 422–438.

[13] Carpanzano, E., Orlandini, A., Valente, A., Cesta, A., Rasconi, R., 2011. “Closed-loop production and automation scheduling in RMSs.” In: Proceedings of the IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). doi: 10.1109/ETFA.2011.6059119.

[14] Uzunosmanoglu, M., Limère, V., Raa, B., Lopez, A.J., Gautama, S., Cottyn, J., Huysentruyt, S., Lamote, J., 2024. “Tactical planning of a reconfigurable cellular assembly system: A case study.” Procedia CIRP 130, 1569–1574. doi: 10.1016/j.procir.2024.10.284.

[15] Albus, M., Hornek, T., Kraus, W., Huber, M.F., 2025. “Towards scalability for resource reconfiguration in robotic assembly line balancing problems using a modified genetic algorithm.” Journal of Intelligent Manufacturing 36 (2), 1175–1199. doi: 10.1007/s10845-023-02292-0.

[16] Padayachee, J., Bright, G., 2012. “Automated design of reconfigurable manufacturing systems: A framework based on a genetic algorithm and discrete system simulation.” In: Proceedings of the 9th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2012), vol. 1, 153–158.

[17] Sisca, F.G., Fiasché, M., Taisch, M., 2015. “A novel hybrid modelling for aggregate production planning in a reconfigurable assembly unit for optoelectronics.” Lecture Notes in Computer Science 9490, 571–582. doi: 10.1007/978-3-319-31972-6_50.

Biography

Elias Huber (*1996) is a researcher at Bosch Corporate Research, focusing on simulation-based planning approaches for reconfigurable robot cells. He received his B.Sc. and M.Sc. in Mechanical Engineering from Karlsruhe Institute of Technology (KIT), Germany. Currently, he is pursuing a Ph.D. at the Institute for Material Handling and Logistics (IFL) at KIT.

Marc Ungen (*1993) is a researcher at Bosch Corporate Research, focusing on simulation-based planning approaches for reconfigurable robot cells. He holds a B.Eng. in Mechanical Engineering and an M.Sc./M.Eng. in Industrial Engineering. Currently, he is pursuing a Ph.D. at the Institute for Control Engineering of Machine Tools and Manufacturing Units (ISW), University of Stuttgart, Germany

8th Conference on Production Systems and Logistics

Developing a Decision-Support Tool for Parameter Configuration in Automated Sorting System

Davide Morchia[1,2], Laura Monferdini[1], Daniele Gibertini[2], Christian Donati[2], Xhulia Sina[3,4], Marco Gotelli[3,4], Eleonora Bottani[1]

[1]*University of Parma, Department of Engineering for Industrial Systems and Technologies (DEIST), viale delle Scienze 181/A, 43124 Parma, Italy*
[2]*CAL SRL - Servizi Logistici, Via Pasubio, 1, 20098 Pasubio MI*
[3]*Simulation Team, Via Magliotto 2, Savona, 17100, Italy*
[4]*DIME, Genoa University, Via Opera Pia 15a, Genova, 16145, Italy*

Abstract

Sorting is one of the logistics processes with the greatest impact on warehouse efficiency and accuracy, as it allows products to be quickly directed to the correct exit points, thus ensuring quick order fulfillment. The objective of this paper is to identify the optimal configuration of the parameters for a real automated sorting system, sited in a company in northern Italy, with the aim of maximizing its performance and productivity. A structured approach, combining various techniques, was followed to achieve the stated goal. As the starting point, a database containing the product batches relating to a year of activity of the company was created. A simulation tool capable of replicating the assignation of batches was developed with VBA under MS Excel™, to replicate the distribution of packages entering the sorter with the logical destination chutes and provide automatically a graphical representation of the batches. The graphs (curves) of the batches were checked for accuracy based on a further data extraction from the sorter management system, referring to the chute assigned to each package, which was compared to the one returned by the simulation tool. The graphs were then gathered into a unique dataset to enable subsequent clustering, so as to classify batches and recognize recurring behaviors. Two different clustering methods were used to this end, based either on statistical parameters (e.g. average chute openings, standard deviation, or variance) or on the curves' data as generated by Excel™. The most performing clustering model returned three clusters, in which the batches were classified; a final regression analysis was made on those clusters to derive a predictive model, able to predict in advance the cluster to which a new batch would belong based on its initial data, thus ensuring its efficient processing. All those elements were embodied into an IT tool, *ad hoc* developed and implemented in the company, to allow the whole procedure to be automated for the computation of the final outcomes. Overall, this study provides the company with an IT tool that: (i) offers objective information (based on real data) on the parameters of the sorting process; (ii) identifies potential areas for improvement of the process itself; and (iii) defines robust criteria for setting efficient operating parameters.

Keywords

Logistics automation; sorting; supply chain performance; clustering; decision-making; parameter optimization.

1. Introduction

Logistics includes all the organizational, management, and strategic processes within a company, from supply to final product distribution [1]. Its purpose is to ensure that the products arrive in the right place and

DOI: https://doi.org/10.15488/20988
ISSN: 2701-6277

at the right time, respecting the established methods, speed, reliability and sustainability [2]. Information management is a key activity in this context, enabling end-to-end traceability, consistent data across systems, and (near) real-time stream visibility [3]. These facilities keep planning and operations aligned, allowing for timely detection of anomalies and the activation of targeted interventions. Accurate and consistent data management reduces traversal times and enables timely and informed decisions to be made across the supply chain [4]. Markets are increasingly dynamic and complex, and supply chains must be agile and flexible enough to quickly adapt to demand, balancing lead time reduction, inventory containment, and on-time delivery [5]. The challenge lies in maintaining accuracy and customer-perceived quality, while keeping costs stable in an ever-changing environment. In this context, the optimization of distribution centers is essential. Distribution centers are central logistics nodes of the modern supply chain, as department stores organized to receive goods from suppliers or other hubs, store them, when necessary, divide them according to destinations and, finally, ship them to points of sale or directly to end customers [6]. Their function is to efficiently connect producers and consumers, ensuring that products arrive on time, in the quantities required, and in the configuration best suited to the destination channel. Unlike a transit point, a distribution center does not simply route goods quickly to the exit docks but also integrates actual storage functions. One of the core activities to ensure efficiency and accuracy of distribution centers is sorting. This process ensures the seamless organization and routing of large volumes of products. A sorter is an automated system that receives parts from the supply lines, identifies them and automatically diverts them to their intended destinations [7]. The sorting operation, and therefore the sorter itself, is governed by a series of parameters that influence the overall capacity and performance of the system. The existing literature focuses on automated sorting systems applied to various sectors and operations. Tan et al. [8] have analyzed the problem of optimizing parcel sorting in an e-commerce warehouse. The authors use a mixed-integer linear programming model and a Particle Swarm Optimisation (PSO) algorithm to solve the problem. Mingrui et al. [9] have also studied automated sorting solutions based on conveyor belts and proposed new modular architectures and advanced programming strategies, including reinforcement learning approaches for high-density parcel sorting. Automated sorting systems include both conveyor-based solutions and emerging robotic sorting systems. Several studies have analyzed robotic sorting architectures as flexible alternatives to traditional conveyor sorters, focusing on layout design, robot congestion, and throughput capacity [10]. The literature on robotic sorting systems has mainly addressed system design, layout selection, and robot assignment policies through analytical and queueing-based models [11]. A recent study [12] has focused on workload balancing and congestion mitigation in robotic sorting systems through dynamic routing and destination reassignment policies (e.g., using MDP and MIP formulations). While previous studies focus on optimization models, robotic systems, or routing policies, they do not explicitly address: (i) the classification of the batch behaviors; (ii) the prediction of batch operational profiles and (iii) their integration into a decision-support tool for parameter setting. This work differs from existing contributions by focusing on decision-support for parameter configuration in real conveyor-based sorting systems, rather than on system design or routing optimization. It proposes a data-driven and implementable framework integrating clustering and predictive modeling to support operational planning. The paper is structured as follows: section 2 describes the context in which the study was conducted and the objective of the paper; section 3 details the methodology, followed by section 4 that discusses the result. Finally, section 5 encompasses the conclusion, limitations and suggestions for future research activities.

2. Context and objective

2.1 Context

This paper examines a real automated sorting system, sited in a company in northern Italy. The company under evaluation owns two circular sorters working in parallel. The packs stored in the receiving area are retrieved by the sorter and moved by conveyor belts. Each sorter consists of a closed ring which moves the

items to the unloading point. Along the perimeter of the ring there are *turrets* (i.e., manual loading stations) that are fed with items. Each sorter has two towers, consisting of 5 infeeders. At each point, operators receive the packages, scan and open them, and place the individual items on the sorter. The items are moved by a launcher (or pusher). Once in the circuit, the items go to the outfeeds (*chutes*) that correspond to a single package designated for a specific store (customer). A store order may be split into several physical packages and therefore into multiple chutes. The number of chutes assigned to a store is determined by the client company according to the total quantity of merchandise required. The focus of the analysis is the *batch*, aimed at creating the packages for different customers. Each batch therefore represents an organized and consistent set of sorting and packaging operations, designed to ensure that the requested products reach the stores in an aggregated, orderly, and traceable form. From an operational standpoint, batches are not generated real time; rather, they are simulated daily by the customer's system. Indeed, the operator must select and define the sequence in which the batches will be executed among those available in the simulated batch portfolio. The screen displayed to the operator contains a series of specific information describing the batch characteristics (Table 1).

Table 1. Summary of the batch-level information displayed to the operator.

Information	Description
Job	Operational activity
12 digits	Simulation date and a sequential number to identify the batches simulated on the same day
date	Day in which the batch simulation has been done
chutes	Number of sorter outputs involved in the batch
packaging	Set of pieces (items)
Pieces	Number of items actually processed
Fill level	Average saturation of the outbound packages
Max chutes	Maximum number of outputs simultaneously occupied by the sorter

The decision to activate the batches is made by the operators according to the following criteria:

- Key performance indicators (KPIs) and timing: the standard Service Level Agreement (SLA) is set at AxB, meaning that batches must be completed within the day after the simulation date.
- Commercial emergencies: urgent batches will follow an AxA SLA.
- Batch structure and resources available for processing.

When the operator chooses to activate a batch on an infeeder turret of the sorter, the entire processing cycle must be completed before proceeding with the next one.

2.2 Objective

The objective of the study is to identify the most effective configurations of the sorter parameters for maximizing its performance and productivity. To this end, the parameters analyzed are:

- The batch activation sequence, i.e., the criteria used to select batches, relating order, and the automatic system on which to process them;
- The replenishment parameter (RP), which defines the quantity of kits to be retrieved from the automatic warehouse to ensure a continuous flow in the sorter.

Overall, this study aims to provide the company with an IT tool that: (i) offers objective information (based on real data) on the parameters of the sorting process; (ii) identifies potential areas for improvement of the process itself; and (iii) embodies robust criteria for setting efficient operating parameters.

3. Methodology

3.1 Data collection

Data was collected on 3,591 batches of products manufactured in one year (May 2024 - July 2025) of activity of the targeted company, to obtain a complete and realistic picture of the workflow. The database includes the batch characteristics described in the previous section.

3.2 Simulation tool

A simulation tool was developed in Microsoft Excel™ using VBA to replicate the assignation of the batches, using both the input and output data extracted from the sorter management system. The input extraction provides, for each batch, the actual temporal sequence of packages arriving at the sorter's infeeder turrets, together with the list of items contained in each package and the relating quantities; the output instead specifies, always for each batch, the item requirements for each logical chute (i.e., which items are needed and in what quantities for completing each chute). By leveraging anticipated knowledge about the arrival sequence of packages and their contents, the VBA macro prevents system saturation and enhances the machine capacity, operating according to the following steps:

i. Reading of the actual sequence of arrival of packages entering the sorter, together with the associated items and quantities contained in each package.
ii. Assigning items to the logical chutes that require them, strictly following the actual chronological order of arrival.
iii. Searching for the optimal distribution configuration that minimizes the opening time of each chute and the number of chutes simultaneously open.
iv. Verifying that each logical chute has received all the item quantities required for its completion and recording the arrival time of each item, defined as the arrival time of the corresponding package at the infeeder plus a deterministic two-minute offset introduced to approximate the internal handling time within the sorter (Eq. 1).

$$t_j^{arr} = t_p + \Delta t \quad (1)$$

Where: t_j^{arr} = arrival timestamp of item j at the logical chute; t_p = arrival time of package p at the infeeder; Δt = deterministic offset equal to 2 minutes.

v. Calculation of the chute opening time, computed as the interval between the arrival of the first and the last required item assigned to that chute (Eq. 2).

$$T_c = \max_{j \in C_c}(t_j^{arr}) - \min_{j \in C_c}(t_j^{arr}) \quad (2)$$

Where: c = logical chute; C_c = set of items assigned to chute c; t_j^{arr} = arrival timestamp of item j at the logical chute

The simulation tool also provides a graphical representation of the temporal progression of the opening and closing of the logical chutes for each batch. The graphs were then gathered into a unique dataset to enable subsequent clustering.

3.3 Clustering for batch classification

To classify batches and recognize recurring behavior types, clustering algorithms were used, reflecting a set of unsupervised learning techniques that ensure the automatic organization of elements into meaningful groups based on their similarities [13]. Two clustering methods were applied, based on (i) statistical parameters (e.g. average chute openings, standard deviation, or variance) and (ii) curves' data as generated by Excel™. For the first approach, a set of statistical and temporal indices was calculated for each curve, including the mean number of open channels, standard deviation, coefficient of variation, skewness, kurtosis, peak-to-mean ratio, relative peak position, interquartile range, Gini index [14], and other measures of

dispersion and centrality. Clustering was performed using the *k*-means algorithm [15], chosen for its efficiency and ability to partition unlabeled observations into homogeneous groups. The optimal number of clusters k was estimated through internal validation metrics, namely the Elbow method and the Silhouette coefficient [16], evaluated over a range of $k = 2$ to 10 clusters Based on these indicators, two candidate solutions ($k = 3$ and $k = 4$) were selected for further analysis. The second clustering method was implemented using the full time-series curves contained in the database. The classification was performed directly on the temporal trajectories describing the opening and closing of the logical chutes for each batch. A first step consisted in normalizing the curves. Since batches differ in both duration and amplitude, direct comparison would introduce distortions in the clustering results. A double normalization procedure was therefore applied: (i) time normalization on the horizontal axis, mapping each batch to a standardized duration; and (ii) amplitude normalization on the vertical axis, scaling the number of open chutes to a common reference range. The optimal number of clusters was determined using the same internal validation indices adopted for the first approach (i.e., Elbow method and Silhouette coefficient), evaluated over a range of candidate k values. Based on this, clustering was generated for $k = 2$, $k = 3$, $k = 4$.

3.4 Outlier removal

The Interquartile Range (IQR) criterion [17] was used to identify anomalous curves, i.e., those exhibiting a substantial different pattern compared to the general behavior of the group and deviating substantially from the centroid of their assigned cluster. The IQR measures the width of the central portion of the distribution, bounded by the first (Q_1) and the third (Q_3) quartiles, and contains 50% of the most frequent values. In accordance with [18], all observations falling below $Q_1 - 1.5 \times IQR$ or above $Q_3 + 1.5 \times IQR$ were considered outliers.

3.5 Predictive regression model

Regression analysis was performed on the clusters obtained from cluster analysis, with the aim of deriving a predictive model capable of forecast the cluster to which a new batch would belong based on its initial data. Through an empirical approach, several regression and classification models were tested and evaluated, including Support Vector Machine (SVM) [19], Random Forest [20], XGBoost [21], k-Nearest Neighbors (KNN) [20], which showed comparable performance in terms of accuracy and precision. These models are widely recognized in the literature as benchmark approaches for classification and prediction tasks in logistics and industrial systems, due to their robustness and ability to capture complex and non-linear relationships. Their inclusion ensures alignment with state-of-the-art predictive modeling practices. In addition to these models, logistic regression was also evaluated, which provided the best performance by translating outputs into class membership probabilities and allowing for direct and robust classification, with a level of interpretability useful in complex operational contexts. For model training and validation, the dataset was partitioned into a training set (80%) and a test set (20%), allowing assessment of the model's ability to generalize to previously unseen data.

3.6 Development of the integrated IT tool

As a final step, an IT tool was developed in Python and implemented in Visual Studio Integrated Development Environment (IDE). The application takes as input the data available for each batch and integrates the logistic regression model developed in the previous step. For each new batch, the model returns the predicted cluster label (1, 2, or 3), corresponding to one of the temporal curve profiles identified by the clustering analysis. This information provides an early estimate of the expected opening/closing dynamics of the logical chutes. A set of decision rules were added to the tool's functions, including: (i) capacity limits on the maximum number of active chutes for each sorter; (ii) a priority scheme for batch execution based on SLA; (iii) a rule linking the current percentage of occupied chutes and the preferred range of kits per outgoing package, used to stabilize the flow; (iv) a set of conditions for combining two batches on the same sorter. The combination is only allowed, depending on the sorter, if: (a) the sum of the respective *max chutes* falls

within predefined safety ranges; (b) sorters have complementary combination (high/low), to balance the load; (c) both batches have ≥ 500 input necks, batches have < 300 chutes in output, and occupied chutes are between 0% and 50% (over 50% the program does not suggest combinations).

The launch time (T_{min}) of a batch and the minimum Replenishment Parameter (RP) were computed in accordance with Eq. (3-4), respectively.

$$T_{min} = \frac{60*K}{\alpha*n*p(\bar{r})} \quad (3)$$

Where: K = number of kits to be entered; n = number of people manning the feeders; $\bar{r}$ = average of swallowing package kits; $p(\bar{r})$ = hourly productivity per operator [kit/hour]

$$RP = \frac{LT}{T \times 60} \times C \times (1 + m) \quad (4)$$

Where: LT = lead time; T = launch duration [minutes]; C = batch neck numbers; m = configurable safety coefficient.

3.7 Experimental set-up for tool validation

A test campaign of five days was conducted to evaluate the effectiveness and reliability of the tool. Specifically, the aim was to verify whether the tool, following its rules, would make decisions similar to the real ones. To build a dedicated validation database, data were gathered from real sorting operations, including the batches activated, the percentage of chutes occupied prior to each activation, the number of active feeders per sorter, the opening and closing times, as well as the monitoring of production breaks. For validation, the software was used to simulate hour by hour the same situation as the 5 days, considering only the batches that, at that time, were available in the real scenario, and assigning each sorter the correct number of operators working that day. Since the simulated batch sequence does not coincide with the actual one, the chute occupancy levels were reconstructed by initializing the first batch with the real value and then applying, for each subsequent batch, the percentage variation observed between consecutive batches in the actual system. The start and completion times of the batches were estimated taking into account the minimum start-up time of the instrument, the operating hours of the plant, and mandatory breaks, with any uncompleted batches carried over to the next day, as in real operations.

4. Results

Figure 1 shows the results of the cluster analysis implemented using the full time-series curves contained in the database and $k = 3$.

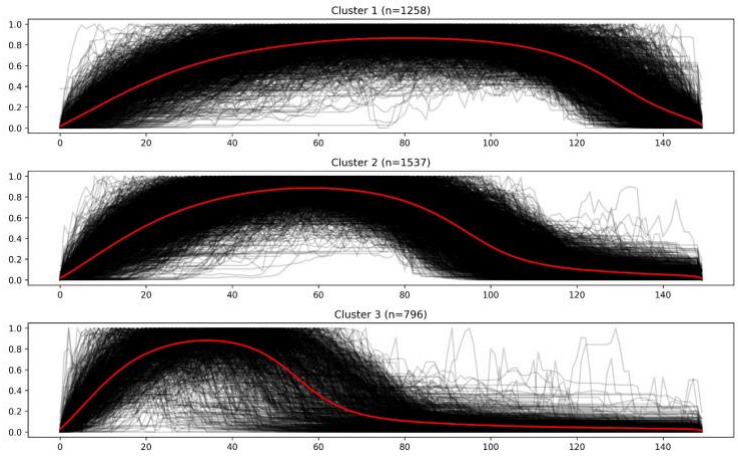


Figure 1. Cluster based on curves with k = 3.

Compared to the cluster based on curves with $k = 2$, which was less well defined, the addition of a third class allows to isolate a group of curves with a trend different from the remaining ones, reducing dispersion and improving the readability of the average trajectories. Although some outliers remain (especially in the third cluster) that do not fit perfectly with the identified classes, the classification is more balanced and capable of better capturing the trend of the curves of the analysed phenomena. Increasing the number of clusters to 4 produces a seemingly more detailed subdivision but which does not translate into an increase in the ability to interpret real batch behaviours. Based on these considerations, the analysis with $k = 3$ offers a sufficiently detailed classification to capture the main differences between the batches.

Figure 2(a-b) shows the results of the logistic regression model and the confusion matrix.

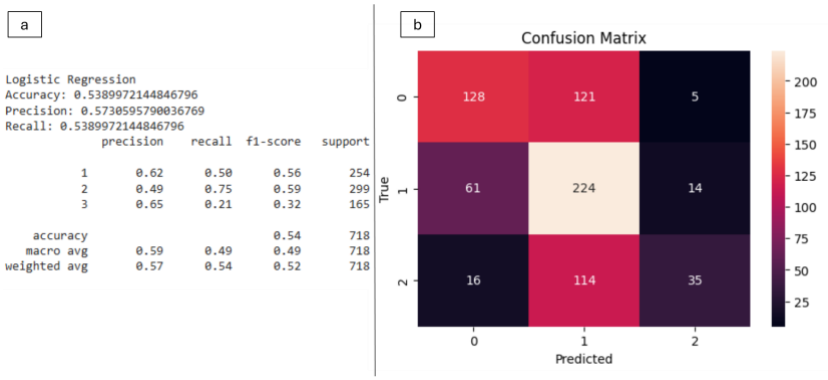


Figure 2. (a) Logistic Regression and (b) Confusion Matrix.

The overall accuracy is 54%, meaning that half of the analyzed batches were correctly assigned to their cluster. The model assigns the batch to the correct cluster in 57% of cases (average accuracy). "Recall" measures the ability to effectively identify all batches belonging to a given cluster; the low value (54%) reflects the model's difficulty in correctly interpreting some observations. The harmonic mean of precision and recall (F1-score) shows differences between clusters: fair results are obtained for cluster 2 (0.59) and cluster 1 (0.56), while cluster 3 is the most problematic with a value of 0.32, indicating that the model has greater difficulty in distinguishing this group from the remaining two. The support parameter indicates the number of observations for each cluster (254, 299, and 165 respectively), useful for understanding the relative weight of each class in training and evaluating the model. The confusion matrix shows that the model captures relevant patterns in batch characteristics with moderate predictive power, particularly for the most represented clusters. Misclassification mainly occur between adjacent clusters, consistent with the continuity observed in the curves during the clustering phase. Model performance is partly limited by the nature of the input variables, which describe structural batch attributes (packages, packaging, filling, etc.), but do not include information from the temporal evolution of the curves, thus reducing the ability to distinguish more complex patterns. Nonetheless, the results suggest that the available indicators are sufficient to reliably predict the batch's operational category for use within the tool's decision-making logic. Table 2 shows the comparison between the real scenario and the one simulated during the five days of testing. The KPI column indicates whether, during the day, all batches were completed within the planned timeline (Yes) or whether one or more batches experienced delays that did not allow meeting the time target (No). The Delays column instead reports the number of batches that, on each day, did not meet the expected timeline, failing to reach the KPI.

Table 2. Comparison of results between real sequence and simulation.

	REAL					SIMULATED				
Day	**Batches Activated**	**Items**	**Combinations**	**KPI**	**Delays**	**Batches Activated**	**Items**	**Combinations**	**KPI**	**Delays**
1	14	123,110	0	Yes	0	15	133,894	0	Yes	0
2	15	157,979	2	No	1	14	146,738	2	Yes	0
3	13	142,460	0	Yes	0	13	142,507	1	Yes	0
4	13	152,122	0	No	2	14	170,418	3	Yes	0
5	9	98,871	1	Yes	0	8	81,085	0	No	1

Total	**64**	**674,642**	**3**	**No**	**3**	**64**	**674,642**	**6**	**No**	**1**

Overall, the total number of batches activated and of packs processed is identical between the two configurations (64 batches and 674,642 packings, respectively), confirming the consistency of the model with respect to the overall observed productivity. The number of batches activated per day, instead, does not perfectly match the simulated sequences, as the model, following its own optimization logic, can actually propose a different activation order, keeping the total work volume unchanged. This behavior, not compromising the result, highlights that the tool is able to reorganize the batch activation based on the simulated operating conditions, while ensuring the complete execution of all expected batches. Regarding the number of batch combinations, 6 combinations were identified overall in the simulated scenario, compared to 3 actually created. This confirms the model's ability to provide a greater number of pairing opportunities between compatible batches. Also, in the current management of the sorting process, delays affected three priority batches, while in the simulated scenario only one batch experienced a delay.

From a numerical point of view, during the testing period:

- The average number of batches activated per day was 12.8, identical for real and simulated sequences, thus providing validation of the tool developed.
- The number of combinations increased by 100% (from 3 to 6), highlighting a better exploitation of plant capacity.
- Delays in batches with active KPI decreased by 66% (from 3 to 1), highlighting better balancing in priority management.
- The overall productivity in terms of packaging processed remained unchanged, confirming that the improvement achieved mainly concerns the quality of planning and not the amount of work performed.

The simulation confirms quantitative consistency with actual data and shows a qualitative improvement in batch pairing capability.

5. Conclusions

This study identified the most effective parameters' configuration for a real automated sorter system, located in a company in Northern Italy. By simulating the current behavior of the plant, an IT tool, able to provide objective information on the process parameters, was developed. This allows for the identification of potential areas for improvement and the definition of robust criteria for setting operating parameters, with the goal of maximizing performance and productivity. The results confirm the robustness of the tool and the consistency of the simulated model with real data, highlighting its ability to accurately represent plant behavior and support task planning. The study highlights some limitations and possible future developments. First, the experimental context was carried out entirely in a simulated environment, to work under controlled conditions, isolating the variables and punctually checking the logical correctness of the model. Although the implemented scheduling logics were found to be robust, a real plant could present greater complexity and slight performance deviations compared to tests could be observed; hence, the real-world applicability of the proposed approach remains unvalidated and testing other real scenarios is recommended. A second limitation concerns the way in which the system's performance parameters are managed. The model calculates how long it takes to launch a batch considering the normal speed at which the operators work on the feeders, the number of operators available and kits that need to be processed, and how "full" the incoming packages are. In real-world contexts, these conditions can be subject to change, which again confirms the appropriateness of further tests in real scenarios; for example, small variations in the rate of feeders or in the number of active workers can alter batch sequencing and overall efficiency. For this reason, it is also advisable to introduce dynamic parameters adjustment, so as to update the tool in real time depending on the observed working conditions, thus enhancing its flexibility. The main opportunity for future development

concerns the integration of the tool with the company's information systems, particularly WMS and ERP. Currently, the collection and processing of data useful to feed the tool are carried out semi-manually; such an approach was adequate for a testing phase, but obviously, it can be improved to fully integrate the system with the production process. As already remarked, further future developments also include validation of the tool in a real production environment. This step will enable the testing of the tool under authentic conditions and the accurate evaluation of its effectiveness in complying with KPIs, reducing delays, and ensuring flow regularity. This phase will also provide an opportunity to gather the feedback of operators and refine interfaces, functionalities and parameters, fostering even greater integration into daily work**.**

References

[1] AILOG, (n.d.). https://www.ailog.it/ (accessed December 1, 2025).

[2] C.J. Langley, M.C. Holcomb, CREATING LOGISTICS CUSTOMER VALUE., Journal of Business Logistics 13 (1992) 1–27. https://openurl.ebsco.com/contentitem/bth:9706191127?sid=ebsco:plink:crawler&id=ebsco:bth:9706191127&crl=c (accessed December 1, 2025).

[3] H.K. Smith, End-to-End Visibility in Entrepreneurial Supply Chains: Leveraging Real-Time Data for Decision-Making, (2024).

[4] O. Aderemi Agboola, A. Chukwuemeke Uzoka, O. Oluwaseun Ajayi, B. Chibunna Ubanadu, A. Ifesinachi Daraojimba, C. Elizabeth Alozie, Oluwademilade Aderemi Agboola Transforming Supply Chain Analytics with Real-Time Data and Cloud Data Warehousing: A Strategic Framework, Engineering and Technology Journal (2025) 2456–3358. https://doi.org/10.47191/etj/v10i05.31.

[5] I.A. Mohammed, J. Mandal, The Impact of Lead Time Variability on Supply Chain Management, International Journal of Supply Chain Management 8 (2023) 41–55. https://doi.org/10.47604/ijscm.3075.

[6] J.K. Higginson, J.H. Bookbinder, Distribution centres in supply chain operations, Logistics Systems: Design and Optimization (2005) 67–91. https://doi.org/10.1007/0-387-24977-X_3.

[7] P. Jung Kunwar, D. Author Kunwar Pukar Jung, S. Karjalainen, V. Pahlsten, Analyzing sorting and packaging for automation and process improvement, (2019). http://www.theseus.fi/handle/10024/267106 (accessed December 1, 2025).

[8] Z. Tan, H. Li, X. He, Optimizing parcel sorting process of vertical sorting system in e-commerce warehouse, Advanced Engineering Informatics 48 (2021) 101279. https://doi.org/10.1016/J.AEI.2021.101279.

[9] M. Yin, H. Zhang, C. Cai, M. Liang, J. Liu, Omniveyor: An Assembled Logistics Sorting System Powered by Reinforcement Learning, IEEE Transactions on Automation Science and Engineering 22 (2025) 11982–11995. https://doi.org/10.1109/TASE.2025.3540511.

[10] B. Zou, R. de Koster, Y. Gong, X. Xu, G. Shen, Robotic Sorting Systems: Performance Estimation and Operating Policies Analysis, Https://Doi.Org/10.1287/Trsc.2021.1053 55 (2021) 1430–1455. https://doi.org/10.1287/TRSC.2021.1053.

[11] N. Boysen, R. de Koster, 50 years of warehousing research—An operations research perspective, Eur. J. Oper. Res. 320 (2025) 449–464. https://doi.org/10.1016/J.EJOR.2024.03.026.

[12] Y. Fang, R. De Koster, D. Roy, Y. Yu, Dynamic Robot Routing and Destination Assignment Policies for Robotic Sorting Systems, Https://Doi.Org/10.1287/Trsc.2023.0458 59 (2025) 603–627. https://doi.org/10.1287/TRSC.2023.0458.

[13] B. Sadeghi, Clustering in geo-data science: Navigating uncertainty to select the most reliable method, Ore Geol. Rev. 181 (2025) 106591. https://doi.org/10.1016/J.OREGEOREV.2025.106591.

[14] F.A. Cowell, Measuring Inequality, (2000). http://sticerd.lse.ac.uk/research/frankweb/MeasuringInequality/index.htmlii (accessed December 10, 2025).

[15] A.K. Jain, Data clustering: 50 years beyond K-means, Pattern Recognit. Lett. 31 (2010) 651–666. https://doi.org/10.1016/J.PATREC.2009.09.011.

[16] J.E. Gentle, L. Kaufman, P.J. Rousseuw, Finding Groups in Data: An Introduction to Cluster Analysis., Biometrics 47 (1991) 788. https://doi.org/10.2307/2532178.

[17] W. Stahel, R.R. Wilcox, Introduction to Robust Estimation and Hypothesis Testing, J. Am. Stat. Assoc. 93 (1998) 1240. https://doi.org/10.2307/2669876.

[18] J.W. Turkey, Exploratory Data Analysis, vol. 2, (1977).

[19] C. Cortes, V. Vapnik, L. Saitta, Support-vector networks, Machine Learning 1995 20:3 20 (1995) 273–297. https://doi.org/10.1007/BF00994018.

[20] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, (2009). https://doi.org/10.1007/978-0-387-84858-7.

[21] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 13-17-August-2016 (2016) 785–794. https://doi.org/10.1145/2939672.2939785;CSUBTYPE:STRING:CONFERENCE.

Biographies

Davide Morchia was born and raised in Piacenza, Emilia-Romagna, in 2001. After graduating as an Electronic and Electrical Engineering Technician from ISII G. Marconi, he continued his studies, graduating in Management Engineering, first with a bachelor's degree and then a master's degree with a specialization in Supply Chain, at the University of Parma. After an internship in logistics, during which the IT tool presented in the article was developed, he joined the company and is currently employed as a logistics engineer at CAL S.r.l.

Laura Monferdini graduated with honors in 2022 in Management Engineering at the University of Parma. She attended the doctoral program in Industrial Engineering with a thesis titled "Frameworks and tOols for evaluating LOgistics processes and enhancing Green, Resilient, Agile and Lean performance (OLOGRAL)". She is currently a research fellow at the University of Parma at the Department of Engineering for Industrial Systems and Technologies.

Daniele Gibertini, born in Tarvisio in 1985, studied at a scientific high school and later graduated in Civil Engineering at the Politecnico di Milano, specialising in Transport Infrastructure. He worked at Expo 2015 from 2011 to 2013 in site mobility and later in Operations. After a brief experience as a traffic consultant, in 2016 he moved into the logistics sector as an engineer at the Technical Office of CAL S.r.l., developing expertise in processes, timing, optimisation and lean management. Since 2020, he has been responsible for the Technical Office of the warehouse described in this publication.

Christian Donati, born in 1994, graduated from the Politecnico di Milano with a degree in Management Engineering and currently works in the Technical Office. Thanks to his previous experience in the sorting department, he was able to collaborate on operational aspects, possessing the analytical skills necessary to support the design, data collection, analysis, and implementation phases of the batch sequencing model.

Xhulia Sina graduated with full marks in Strategic Engineering from the University of Genoa and holds a bachelor's degree in Mechanical Engineering. She is currently pursuing an industrial PhD as a Sim4Future Data Scientist at the University of Genoa, within the DIME Department. She is a member of the Simulation Team, collaborating on projects with international organizations and industrial partners. She has a strong background in modeling and simulation, focusing on applying ML to optimize industrial systems.

Marco Gotelli is Researcher at the University of Genoa, where he earned his Master's degree in Mechanical Engineering and a PhD in Mathematical Modeling and Simulation. His research focuses on modelling and simulation, digital transformation and data-driven risk assessment for industrial, logistic and port systems, building on prior experience as Senior Consultant at Deloitte and on research activities at the Fraunhofer Institut

Eleonora Bottani is Full Professor at the University of Parma, where she graduated with honors in 2002 in Management Engineering and got a PhD in Industrial Engineering in 2006. Her research focuses on logistics and supply chain and have been published in more than 270 journal and conference papers (citations on Scopus >5600; H-index=37).

8th Conference on Production Systems and Logistics

Bridging the Synchrony Gap: A Deterministic GSMDP Approach for Integrating Deep Reinforcement Learning with Commercial Discrete Event Simulators

Merve Demir Varici, Hasan Tercan, Tobias Meisen
Institute for Technologies and Management of Digital Transformation der Bergischen Universität Wuppertal , Wuppertal, Germany

Abstract

Although Deep Reinforcement Learning (DRL) offers adaptive control for manufacturing processes, training agents within high-fidelity commercial Discrete Event Simulation (DES) tools is hindered by the temporal mismatch between the continuous-time event logic of DES and the discrete-step nature of RL. Current loose-coupling methods (e.g., socket communication) frequently introduce latency and state drift, violating the Markov assumption necessary for stable training. This study introduces a strictly blocking, in-memory synchronization architecture that aligns DES execution with Generalized Semi-Markov Decision Process (GSMDP) decision epochs. Unlike asynchronous approaches, our framework freezes the simulation clock at decision points to ensure near-zero-latency state observations. We validated the architecture using a Dual-Resource Constrained Job Shop Scheduling Problem (DRC-JSSP) involving sequence-dependent setup times. Evaluations on this test bed demonstrate that the agent can learn near-optimal policies with a deviation of only 2.24% from the optimal CP-SAT solver, even under complex constraints where synchronization errors would otherwise lead to infeasible schedules.

Keywords

Deep Reinforcement Learning; Discrete Event Simulation; GSMDP Synchronization; Scheduling; Dual-Resource Constrained Job Shop Scheduling

1. Introduction

In the Industry 4.0 era, simulation platforms have evolved from offline analytical tools into the backbone of digital twins, facilitating real-time decision-making in complex stochastic systems [1]. Although traditional heuristic rules are commonly employed for scheduling purposes, they frequently struggle to adapt to dynamic disturbances, such as machine breakdowns or urgent orders. This limitation has driven a shift towards data-driven control methods, particularly Deep Reinforcement Learning (DRL), which can learn adaptive policies directly through interaction with a simulation environment [2] [3].

The concept of integrating simulation with learning-based decision-making is not new, having been explored for decades in operations research [4], [5], [6]. However, the field has been revitalized in recent years due to the emergence of DRL, which enables the handling of high-dimensional state spaces that were previously computationally intractable [7], [8].

Despite the potential of DRL, there is still a significant technical gap when it comes to deploying modern algorithms within commercial Discrete Event Simulation (DES) platforms such as AnyLogic and Siemens

DOI: https://doi.org/10.15488/20989
ISSN: 2701-6277

Plant Simulation. Academic research predominantly relies on lightweight, custom-built Python environments (e.g., OpenAI Gym, simplistic grid worlds or specific libraries such as ProdSim) to facilitate rapid integration and training [9], [10]. While flexible, these custom environments often lack the visual fidelity, validation capabilities and complex logic libraries provided by established commercial platforms such as AnyLogic, Siemens Plant Simulation or Arena, which are increasingly being adopted in industrial digital twin applications [9]. Conversely, integrating Python-based agents with these commercial simulators typically relies on high-latency communication protocols (e.g., TCP/IP sockets), which introduce synchronization errors and significantly slow down the training process [11].

Beyond basic software interfacing, the fundamental challenge of this integration lies in the temporal misalignment between the event-driven progression of DES and the discrete decision steps of RL. This interaction can be described mathematically as a Generalized Semi-Markov Decision Process (GSMDP), where the time between decision epochs is stochastic and state-dependent [12]. Standard asynchronous integrations often violate the Markov assumption by allowing the simulation clock to advance while the agent computes an action. To address this issue, we propose a strictly blocking in-memory architecture. Unlike loose-coupling methods, our framework freezes the simulation clock at GSMDP decision epochs, ensuring near-zero-latency observations and preserving the mathematical integrity of the state transition $S_t \rightarrow S_{t+1}$.

This work proposes a deterministic, strictly blocking in-memory integration architecture approach as a proof-of-concept for coupling DRL agents with commercial DES platforms, explicitly addressing the temporal misalignment and latency effects inherent in asynchronous integration approaches reported in the literature [11]. We evaluate the architecture on a Dual-Resource Constrained Job Shop Scheduling Problem (DRC-JSSP) to verify its ability to maintain temporal consistency at GSMDP decision epochs. By leveraging a high-fidelity simulation model, this study illustrates how such an integration can support more realistic training settings and reduce the sim-to-real gap commonly observed in simplified RL scheduling environments [13].

The remainder of this paper is organized as follows: Section 2 provides a review of the relevant background to RL, DES, and existing integration approaches. Section 3 discusses the challenge of temporal synchronization. Section 4 presents the strictly blocking in-memory integration architecture proposed in this study. Section 5 evaluates the framework using the DRC-JSSP test bed. Section 6 concludes the paper and outlines future work.

2. Theoretical Background and Related Work

This section reviews the intersection of DES and RL, focusing specifically on integration architectures and the theoretical challenges of temporal synchronization in industrial environments.

2.1 RL Integration in Commercial Simulation Platforms

Although RL has shown great promise in production scheduling, the significant sim-to-real gap remains a barrier to industrial adoption [14]. Academic research predominantly uses lightweight, custom-built Python environments (e.g., Gym wrappers) to train agents. However, these environments often lack the fidelity required to model complex shop-floor dynamics, such as material handling, worker shifts, and stochastic breakdowns [11]. Such operational constraints are repeatedly identified in the literature as key determinants of industrial realism, which are often overlooked in standard RL benchmarks [10], [15], [16].

Conversely, commercial DES platforms (e.g., AnyLogic, Siemens Plant Simulation) offer high-fidelity modeling but lack native, high-frequency interfaces for RL control. A major barrier to using these robust tools for RL experimentation is the absence of standardized, bidirectional, high-frequency communication interfaces with external learning agents [9] [15].

Recent attempts to bridge this gap have typically relied on external communication protocols. For example, Hubert et al. [8] implemented a TCP/IP socket-based interface to connect a Python agent with a DES model. Similarly, Peyman et al. [11] demonstrated connector architectures for various commercial tools. However, these loosely coupled architectures introduce communication latency and serialization overhead, which can decouple the simulation clock from the agent's decision loop, ultimately leading to non-Markovian state observations.

Another strategy reported in the literature is epoch-based learning, in which the agent only updates its policy after completing a simulation run. Although this avoids real-time synchronization conflicts, it is highly sample-inefficient and prevents the agent from reacting to intra-episode dynamics or learning from immediate state transitions [15].

2.2 The Synchrony Gap: From MDP to GSMDP

Integrating RL with DES poses a fundamental challenge in the form of a temporal mismatch. Standard RL algorithms assume a discrete-time Markov decision process (MDP), in which interactions occur at fixed time steps. In contrast, DES operates in continuous time and advances at irregular intervals, which are triggered by events (e.g., machine idling or job arrival) [12].

This continuous-discrete interaction can be formally described as a GSMDP, in which the time between decision epochs is stochastic and depends on the system state [12]. In loose-coupling approaches (e.g., sockets), the simulation clock (τ) often continues to advance while the agent computes an action (i.e., during inference time). This creates state drift, whereby the action is applied to state $S_{\tau+\Delta t}$ rather than the observed state S_{τ}. Our work addresses this issue by enforcing a strictly blocking synchronization mechanism that ensures the simulation engine freezes deterministically at every decision epoch, thus preserving the mathematical integrity of the GSMDP transition.

2.3 Dual-Resource Constrained Job Shop Scheduling

To validate the proposed architecture, we address the DRC-JSSP. Unlike the classical JSSP, which considers only machine availability, the DRC-JSSP incorporates secondary constraints, such as the availability of skilled workers or auxiliary tools. This problem class is NP-hard and particularly sensitive to synchronization errors. For instance, in a DRC environment, an operation can only commence when a machine and a specific tool are both available simultaneously. If the RL agent's decision is delayed due to communication latency, as occurs with socket-based methods, another process may seize the available tool during the delay, rendering the agent's action invalid. Therefore, the DRC-JSSP problem class is a controlled testbed to verify the precision of our near-zero-latency integration framework compared to traditional heuristic rules or loosely coupled methods [17].

3. Proposed Framework: A GSMDP-Based Synchronization Architecture

3.1 The Necessity of Blocking Synchronization

Ideally, an RL agent interacts with a simulation in a strict sequence: Observation (S_t) → Action (A_t) → Reward (R_t) → Next State (S_{t+1}). However, commercial DES platforms like AnyLogic operate on a continuous timeline. Existing integration methods often use TCP/IP sockets for communication [9]. While flexible, socket-based communication introduces non-deterministic latency (communication overhead). In loosely coupled architectures, the simulation clock (τ) continues to advance during the agent's inference time (Δt_{inf}) [8], [18].

In high-speed training, this latency causes state drift. Consequently, the action A_t is applied to a drifted state $S_{\tau} + \Delta$, violating the Markov property required for stable convergence. This temporal misalignment breaks the Markov assumption because the state transition becomes dependent on the stochastic variable Δt_{inf}, an

artifact of the communication architecture rather than the true system dynamics [14], [19]. To solve this, our framework enforces a strictly blocking synchronization mechanism, as illustrated in Figure 1. The simulation engine is algorithmically suspended (frozen) at the exact moment of a decision epoch, and remains frozen throughout the agent's inference phase, ensuring that the inference time of the agent (Δt_{inf}) does not affect the simulation clock, which remains constant with d_τ / d_t equal to zero during inference.

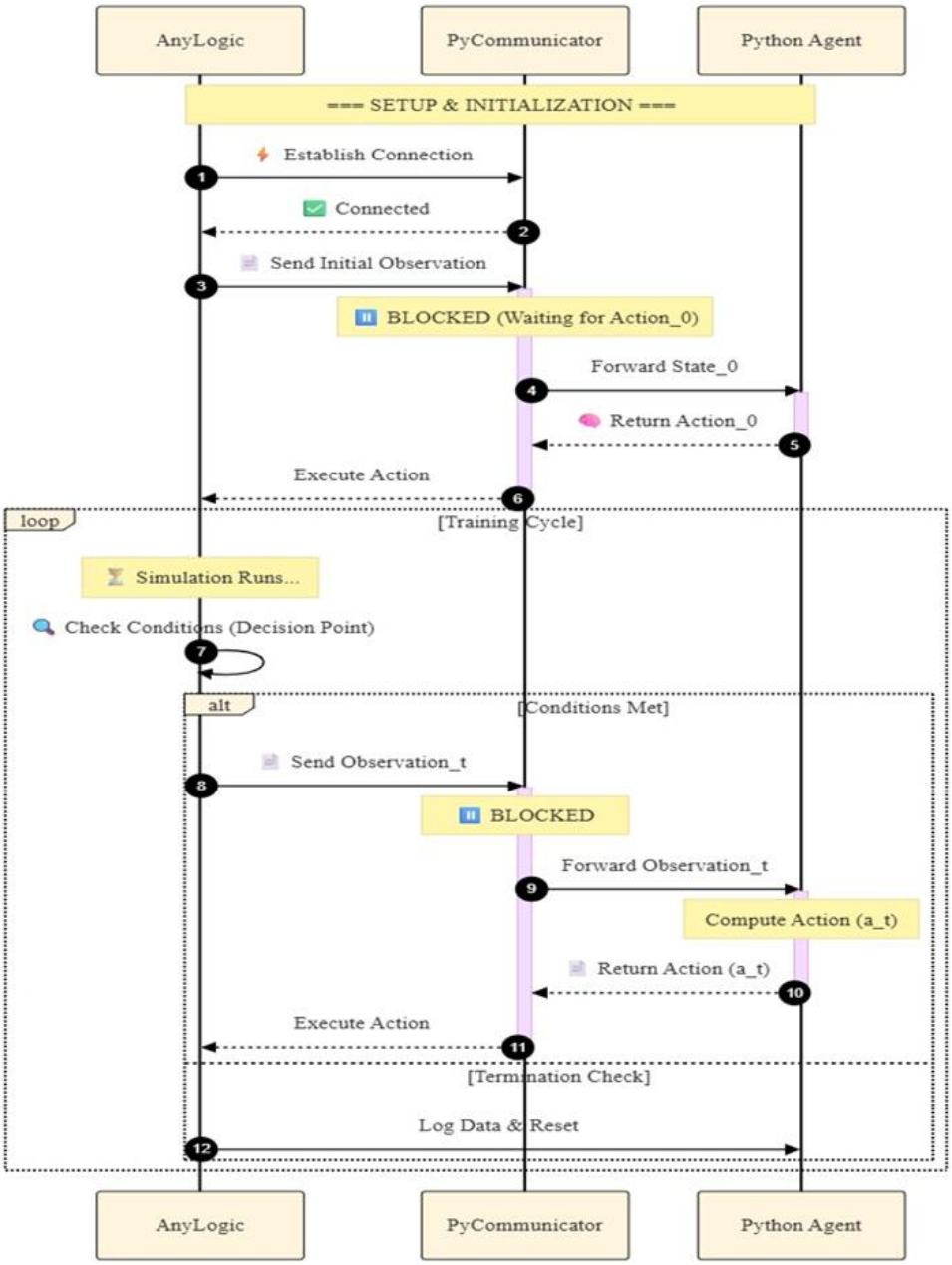


Figure 1: Structural illustration of the agent–environment communication loop and the essential dynamics within the environment.

3.2 In-Memory Architecture via Pypeline

To implement this blocking mechanism without incurring the serialization overhead associated with external sockets, we utilize an in-memory bridge architecture. Unlike the distributed client-server models reviewed in [9], our approach embeds the Python interpreter directly within the AnyLogic Java Virtual Machine (JVM). This architecture offers two key benefits: (1) near-zero-latency data exchange, whereby state vectors are passed as memory references rather than serialized JSON/XML packets, thus eliminating the I/O

bottleneck described by Hubert et al. [8]; and (2) strict thread synchronization. When the DES model triggers a decision request, the Java execution thread invokes the Python agent and enters a wait state. The thread only resumes after the Python function returns the action integer, thus ensuring that no simulation events can occur during the decision-making process.

3.3 Event-Triggered Decision Epochs

In accordance with the GSMDP formalism [12], the decision epochs in our framework are stochastic, state-dependent events rather than fixed time steps. Although event-triggered control is well-established in the Semi-Markov Decision Process (SMDP) literature [20], implementing such synchronization within commercial simulators introduces practical latency and execution-control challenges.

To manage these challenges, our architecture explicitly defines the interaction moments. Specifically, a decision point is triggered only when the following feasibility conditions are met simultaneously:

- (C1) A machine becomes idle.
- (C2) The queue for that machine is not empty.
- (C3) The required auxiliary resources (e.g., tools, workers) are available.

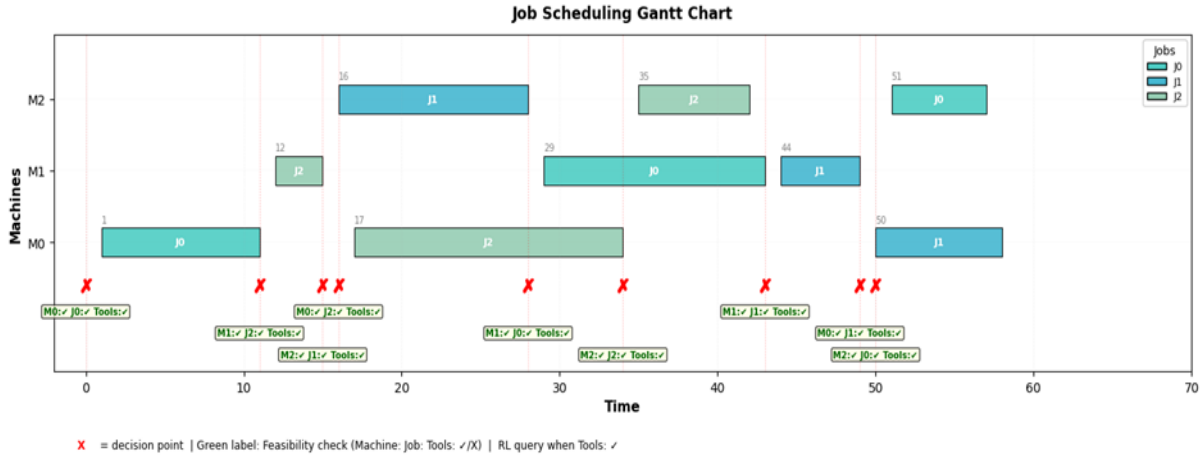


Figure 2: Gantt chart illustrating decision points (red) and feasibility checks (green) used to trigger interactions between the simulation and the RL agent.

As illustrated in Figure 2, if a machine becomes idle but the required tool is occupied by another machine, the decision point is suppressed due to the dual-resource constraint. This selective triggering prevents the agent from selecting invalid actions, thereby reducing the complexity of the action space and accelerating convergence.

3.4 State Representation and Reward Formulation

The proposed framework is architecturally designed to support dense, stepwise feedback and detailed state representations at every decision epoch, enabling complex reward shaping strategies. Following each action, a near-zero-latency snapshot of the shop-floor dynamics is transferred to the agent to inform the subsequent decision. However, to systematically validate the stability of the synchronization mechanism and the agent's ability to solve the combinatorial optimization problem without heuristic guidance, this study employs a sparse terminal reward formulation.

The reward is calculated only upon the termination of an episode, when all jobs are completed. The reward function is defined as the negative makespan, $R = -C_{max}$. This formulation encourages the agent to learn the global dependency between early scheduling decisions and final production efficiency, rather than prioritizing short-term greedy gains. The ability of the Pypeline architecture to effectively train an agent under these sparse reward conditions provides empirical support for the temporal consistency and

synchronization reliability of the integration, as synchronization errors would disrupt the causal relationship between actions and the delayed reward.

4. Experimental Validation: Dual-Resource Constrained Scheduling

To evaluate the fidelity of the proposed GSMDP synchronization architecture, we apply the framework to the DRC-JSSP. While many RL studies focus on large-scale problems with simplified constraints, this study intentionally utilizes a compact, highly constrained 3×3 instance as a verification benchmark. This controlled setting enables us to compare the RL agent's learning trajectory with a theoretically optimal solution calculated using CP-SAT. This comparison verifies that the blocking synchronization mechanism mitigates the state drift typically observed in larger, loosely coupled simulations.

4.1 Problem Specification

The experimental use case models a dual-resource constrained (DRC) manufacturing cell comprising three jobs (J1, J2, and J3), which are processed across three machines (M1, M2, and M3) with distinct routing sequences. Unlike standard JSSP formulations, this environment imposes strict, coupled resource constraints; operations require the simultaneous availability of the primary machine and specific, finite-capacity auxiliary tools (T1–T6), as detailed in Table 1. Furthermore, the model incorporates sequence-dependent setup times, whereby the preparation of a machine depends on the job that preceded it. The combination of these synchronization bottlenecks and nonlinear temporal dependencies creates a testing environment to evaluate the agent's ability to learn long-term scheduling policies under tightly constrained conditions.

Table 1: Tool requirements and capacities in the proposed DRC-JSSP model.

	Machine 0	**Machine 1**	**Machine 2**
Job 0	T1; T4; T5	T1; T5	T1; T6
Job 1	T2; T4; T6	T2; T5	T2; T6
Job 2	T3; T4; T6	T3; T5	T3; T4
Tool capacities:	T1 ( ):1, T2 ( ):1, T3( ):2, T4 ( ):1, T5( ):1, T6 ( ):1		

4.2 RL Agent Configuration

The RL process is executed using the Proximal Policy Optimization (PPO) algorithm from the Stable-Baselines3 library, integrated via the Pypeline interface.

• Observation Space (S_t): The state is represented as a flattened vector containing the availability of auxiliary tools alongside five key operational parameters for each machine k: (1) the real-time occupancy status $\delta_k \in \{0,1\}$, (2) the number of remaining tasks in the queue, (3) the ID of the next machine in the job's routing sequence, (4) the expected processing time required at that next machine, and (5) the remaining processing time of the current operation (τ_k^{rem}). Crucially, thanks to the blocking architecture, this observation S_t is guaranteed to be a near-zero-latency snapshot of the simulation at the exact decision epoch.

• Action Space (A_t): This is a discrete action space representing the selection of a waiting job from the queue of the idle machine.

• Reward Function (R_t): To rigorously test the agent's ability to solve the credit assignment problem without heuristic guidance, we utilize a sparse terminal reward: $R = -C_{max}$. The reward is calculated only at the end of the episode (negative makespan). Successfully learning from such a sparse signal serves as a validation of the architecture's temporal consistency. If there were synchronization errors or latency between

the Python agent and AnyLogic, the causal link between early actions and the final makespan would be broken, preventing convergence.

4.3 Results and Benchmarking

The agent was trained over 10,000 episodes in headless mode to accelerate data collection. To validate the quality of the learned policy, the results were compared with the theoretical optimum, representing a lower bound obtained using the Google OR-Tools CP-SAT solver under identically modeled constraints, as well as with standard heuristic dispatching rules, namely shortest processing time (SPT) and first in, first out (FIFO).

Table 2: Performance comparison (Makespan) on test instances.

Method	Case A Makespan	Case B Makespan	Case C Makespan	APD
Optimum	53	42	58	Reference Value
PPO Agent	57	43	58	2.24%
SPT Rule	78	50	93	49.65%
FIFO Rule	65	43	70	21.46%

APD: Average Percentage Deviation from the theoretical optimum.

4.4 Discussion of Results

As shown in Table 2, the RL agent trained using the proposed architecture achieves a makespan that differs from the CP-SAT optimum by only 2.24% on average.

Validation of Synchronization: The agent's significant outperformance over heuristic rules (SPT: 49.65% gap; FIFO: 21.46% gap) and its near-optimal approach to the CP-SAT solution provide strong empirical evidence that Pypeline-based blocking synchronization preserves the Markov property. The agent was able to learn complex dependencies (e.g., waiting for a specific tool to become available), which are invisible to myopic heuristics.

Architectural Implications: Although CP-SAT finds the optimal solution faster for this specific static instance, the RL agent's near-optimal performance confirms that the simulation-interaction loop is mathematically sound. Unlike the exact solver, which requires a pre-defined mathematical model, the RL agent learned the constraints purely through interaction. While the evaluation is currently limited to a small-scale deterministic benchmark, this proof-of-concept demonstrates the technical feasibility of the synchronization mechanism, suggesting potential applicability for more complex, stochastic scenarios (e.g., machine breakdowns) where static solvers such as CP-SAT are not directly applicable.

5. Conclusion and Future Work

This study bridges the technological gap between high-fidelity commercial DES platforms and DRL algorithms. While Python-based environments are common in research, they often lack the industrial realism required for Digital Twins. Conversely, standard socket-based integrations with tools such as AnyLogic suffer from latency and temporal misalignment. To address this, we formalized and validated a strictly blocking in-memory architecture that enforces GSMDP-compliant synchronization, ensuring near-zero-latency observations from the perspective of the simulation clock and preventing state drift during training.

The efficacy of this framework was verified using the DRC-JSSP under a sparse reward setting. The PPO agent successfully learned a policy with only a 2.24% deviation from the optimal CP-SAT solution, demonstrating that the architecture maintains the Markov property even under complex constraints such as sequence-dependent setups and tool availability. This confirms that the specific synchronization

implementation effectively preserves the mathematical integrity of the learning process within the tested boundaries.

Future work will evaluate the scalability of this integration approach to larger industrial problems and address dynamic stochasticity, such as random machine breakdowns and urgent order insertions—scenarios where static solvers typically fail but where RL agents can demonstrate superior adaptability. Additionally, we aim to explore multi-agent configurations to further enhance responsiveness in large-scale manufacturing systems.

Acknowledgements

The first author, Merve Demir Varici, acknowledges the financial support provided by the Study Abroad Postgraduate Education Scholarship (YLSY) program of the Republic of Türkiye Ministry of National Education.

References

[1] Qiu, H., Chen, Y., Zhang, H., Yi, W., Li, Y., 2022. Evolutionary Digital Twin model with agent-based discrete-event simulation method. In: Proceedings of the 8th Conference on Production Systems and Logistics. Springer, pp. 6178–6183.

[2] Panzer, M., Bender, B., 2022. Deep reinforcement learning in production systems: A systematic literature review. International Journal of Production Research 60 (13), 4316–4341.

[3] Lv, L., Zhang, C., Fan, J., Shen, W., 2025. Deep reinforcement learning for job shop scheduling problems: A comprehensive literature review. Knowledge-Based Systems 321.

[4] Zeng, Q., Yang, Z., Hu, X., 2011. A method integrating simulation and reinforcement learning for operation scheduling in container terminals. Transport Reviews.

[5] Allaoui, H., Artiba, A., 2004. Integrating simulation and optimization to schedule a hybrid flow shop with maintenance constraints. Computers & Industrial Engineering 47 (4), 431–450.

[6] Aydin, M., Öztemel, E., 2000. Dynamic job-shop scheduling using reinforcement learning agents. Robotics and Autonomous Systems 33 (2–3), 169–178.

[7] Pilarski, S., Sidhu, A., Varró, D., 2023. A simulation environment for reducing food waste via reinforcement learning. In: Proceedings of the Annual Modeling and Simulation Conference.

[8] Hubert, S., Meintschel, J., Bleidorn, D., Ortmanns, Y., Wallrath, R., 2023. Production scheduling using deep reinforcement learning and discrete event simulation. Chemie Ingenieur Technik 95 (1003–1011).

[9] Kemmerling, M., Samsonov, V., Lütticke, D., Schuh, G., Gützlaff, A., Schmidhuber, M., Janke, T. K., Meisen, T., 2021. Towards production-ready reinforcement learning scheduling agents: A hybrid two-step training approach based on discrete-event simulations. In: Simulation in Produktion und Logistik, Göttingen, pp. 1–12.

[10] Fuchs, T., Enslin, C., Samsonov, V., Lütticke, D., Schmitt, R. H., 2022. ProdSim: An open-source Python package for generating high-resolution synthetic manufacturing data on product, machine and shop-floor levels. Procedia CIRP 104, 1343–1348.

[11] Peyman, M., Copado, P., Panadero, J., Juan, A. A., Dehghanimohammadabadi, M., 2021. A tutorial on how to connect Python with different simulation software to develop rich simheuristics. In: Proceedings of the IEEE International Conference on Industrial Informatics.

[12] Capocchi, L., Santucci, J.–F., 2022. Discrete event modeling and simulation for reinforcement learning system design. Information 13, 420.

[13] Liang, H., Xie, J., Huang, B., Li, Y., Sun, B., Yang, C., 2025. A novel sim2real reinforcement learning algorithm for process control. Reliability Engineering & System Safety 254.

[14] Da, L., Turnau, J., Kutralingam, T. P., Velasquez, A., Shakarian, P., Wei, H., 2025. A survey of sim-to-real methods in RL: Progress, prospects and challenges with foundation models. arXiv preprint arXiv:2502.13187.

[15] Grzona, P., Yen, M. T., Zumpe, F., Münnich, M., 2023. Requirements for a ML- and platform-based simulation service. Simulation in Produktion und Logistik, pp. 187–195.

[16] Zhou, Y., Du, S., Liu, M., Shen, X., 2024. Machine-fixture-pallet resource constrained flexible job shop scheduling considering loading and unloading times under pallet automation system. Journal of Manufacturing Systems 73, 143–158.

[17] Liang, G., Quan-Ke, P., 2016. A shuffled multi-swarm micro-migrating birds optimizer for a multi-resource constrained flexible job shop scheduling problem. Information Sciences 372, 655–676.

[18] Samsonov, V., Kemmerling, M., Paegert, M., Lütticke, D., Sauermann, F., Gützlaff, A., Schuh, G., Meisen, T., 2021. Manufacturing control in job shop environments with reinforcement learning. In: Science and Technology Publications.

[19] Derman, E., Dalal, G., Mannor, S., 2021. Acting in delayed environments with non-stationary Markov policies. arXiv preprint arXiv:2101.11992.

[20] Mahadevan, S., Marchalleck, N., Das, T. K., Gosavi, A., 1997. Self-improving factory simulation using continuous-time average-reward reinforcement learning. In: Proceedings of the Morgan Kaufmann Publishers Conference on Simulation.

Biography

Merve Demir Varici (*1994) holds a master's degree from Erzurum Ataturk University in Industrial Engineering. She received a scholarship from the Türkiye Ministry of National Education for her doctoral study. Since 2022, she works at the Institute for Digital Transformation Technologies and Management at the University of Wuppertal. Her research focuses on solving planning and scheduling problems with deep reinforcement learning.

Hasan Tercan (*1988) holds a master’s degree from TU Darmstadt in Computer Science. Since 2018, he is a scientific researcher at the Institute for Technologies and Management of Digital Transformation and leader of the research group Industrial Deep Learning. In his research, he investigates the application of machine learning methods in industrial processes.

Tobias Meisen (*1981) is a professor for Technologies and Management of Digital Transformation at the University of Wuppertal since 2018. He is also the institute director of the In-Institute for Systems Research in Information, Communication and Media Technology, the chair of the Interdisciplinary Centre for Data Analytics and Machine Learning.

8th Conference on Production Systems and Logistics

Logistics Model To Reduce Inventory Discrepancies By Applying Digital POKA-YOKE, RFID And WMS In The Warehouses of White-Goods Company

Carlos Ali Irigoin-Sánchez, Diego Vincenzo Mascaro-Levano,
Jorge Antonio Corzo-Chavez, Gino Viacava-Campos
Carrera de Ingeniería Industrial, Universidad de Lima, Lima, Perú

Abstract

Inventory accuracy is a key element for maintaining reliable and efficient logistics operations. In this context, this study analyzes record accuracy and proposes innovative digital solutions. This study addresses the problem of inventory record accuracy (IRA) in white-goods warehouses, where a huge discrepancy was detected between the physical inventory and the system records, resulting in an accuracy level of 89%, below the industry benchmark of 98%. This situation generates annual economic losses of USD 129,888, rework, and low reliability in logistics decision-making. An integrated improvement model is proposed, based on three tools: Pick-to-Light, aimed at reducing picking errors through visual guidance; RFID with overload alerts, designed for real-time physical validation and traceability; and a digital Poka-Yoke, implemented to automatically block incorrect or incomplete records. The study demonstrates that the integration of digital control and traceability tools represents an innovation applicable to logistics management, combining Lean technologies with intelligent validation systems to address inventory accuracy from a preventive and automated perspective. This approach goes beyond traditional physical audits, proposing a layered control model that ensures real-time coherence between physical and digital inventory. Its modular and adaptable structure allows the proposal to be replicated in other industrial sectors or logistics contexts, where traceability and inventory accuracy are critical factors for supply chain efficiency. Overall, the study contributes a validated and replicable methodological framework that advances knowledge in digital logistics management and continuous improvement. Validation was carried out through computational simulation, comparing the current and improved scenarios, revealing an increase in inventory accuracy from 89% to 98%, a significant decrease in picking, dispatch, and operational errors, and annual savings exceeding USD 75,000. The application of digital and control tools under a Lean-logistics approach constitutes an effective alternative to strengthen traceability, security, and inventory reliability in the white-goods sector.

Keywords

Inventory accuracy; RFID; Pick-to-Light; Digital Poka-Yoke; Lean logistics

1. Introduction

1.1. Background

Inventory record accuracy (IRA) is a critical issue in white-goods warehouses, where high product turnover, multiple stock-keeping units, and reliance on manual processes increase the likelihood of discrepancies [1].

DOI: https://doi.org/10.15488/20990
ISSN: 2701-6277

In the analyzed case, an inventory accuracy level of 89% was identified, significantly below the industry benchmark of 98%, generating estimated annual losses of USD 129,888, along with rework, dispatch delays, and reduced reliability in logistics decision-making [2], [3].

Operational analysis revealed recurrent discrepancies between physical inventory and ERP system records, mainly caused by errors in picking, dispatch validation, and administrative transactions in SAP (MIGO) [1]. However, existing solutions are typically implemented as isolated technologies or sequential validation mechanisms, where discrepancies are detected after operations have been completed rather than prevented in real time. In addition, the absence of real-time validation mechanisms and traceability systems contributes to increased operational variability and human error [4]. Although previous studies highlight the benefits of technologies such as RFID, Pick-to-Light, and Poka-Yoke in improving inventory accuracy [2], [4], [5], limited research has addressed their integration within control architectures that ensure real-time synchronization between physical operations and ERP-based inventory records [3].

To address this gap, this study proposes a Lean–Digital control framework that integrates Pick-to-Light systems, RFID-based validation with overload alerts, and a digital Poka-Yoke embedded in the ERP system to prevent inconsistent transactions. The model was evaluated using discrete-event simulation based on real operational data. The results indicate that IRA can be improved from 89% to approximately 98% under simulated conditions, while significantly reducing picking errors, missing units, and transactional discrepancies. The proposed framework also suggests potential annual savings exceeding USD 75,000. This study contributes by formalizing a layered control architecture that enforces real-time consistency between physical and digital inventory records, thereby shifting inventory management from error detection to error prevention in warehouse operations.

1.2. Literature Review

The literature review followed a PRISMA-based protocol to ensure methodological rigor, transparency, and reproducibility. Searches were conducted in the Scopus and Web of Science databases using the keywords “inventory record accuracy,” “RFID warehouse,” “pick-to-light,” “digital poka-yoke,” and “warehouse management systems” for the period 2015–2024. A total of 469 records were initially identified (Scopus = 102; Web of Science = 367). After removing 163 duplicate records, 306 articles remained for title and abstract screening. During this stage, 214 studies were excluded due to lack of relevance to warehouse operations, absence of focus on inventory accuracy, or unrelated technological scope. Consequently, 92 articles were selected for full-text eligibility assessment. At the eligibility stage, 66 full-text articles were excluded for the following reasons: lack of direct focus on inventory record accuracy (n = 24), absence of quantitative or empirical validation (n = 18), purely conceptual or review-based orientation without operational evidence (n = 14), and insufficient relation to IRA or digital validation technologies such as RFID, PTL, or digital Poka-Yoke (n = 10). Finally, 26 peer-reviewed journal articles met all inclusion criteria and were included in the qualitative synthesis. Inclusion criteria comprised peer-reviewed journal articles written in English, focused on warehouse operations or inventory accuracy, and providing quantitative or empirical evidence.

The state of the art, summarized in Figure 1, highlights the critical role of digital technologies in modern supply chain management as fundamental mechanisms for mitigating operational errors and improving inventory reliability. Specifically addressing Inventory Record Accuracy (IRA), the literature shows that RFID technology integrated with Warehouse Management Systems (WMS) enhances product traceability and real-time visibility, significantly reducing discrepancies between physical and digital inventories [3]. Similarly, Pick-to-Light (PTL) systems are recognized as effective visual Poka-Yoke mechanisms that reduce process variability and minimize operator search times during order preparation [5]. Consequently, prior studies consistently demonstrate that integrating Lean tools with digital validation systems produces

measurable improvements in productivity, operational control, and inventory accuracy [6]. However, existing studies predominantly analyze these technologies as isolated solutions or sequential validation mechanisms within warehouse operations. Limited attention has been given to integrated control architectures that enforce real-time synchronization between physical execution and digital inventory records. As a result, most approaches remain focused on error detection rather than prevention, allowing discrepancies to propagate across operational and transactional stages.

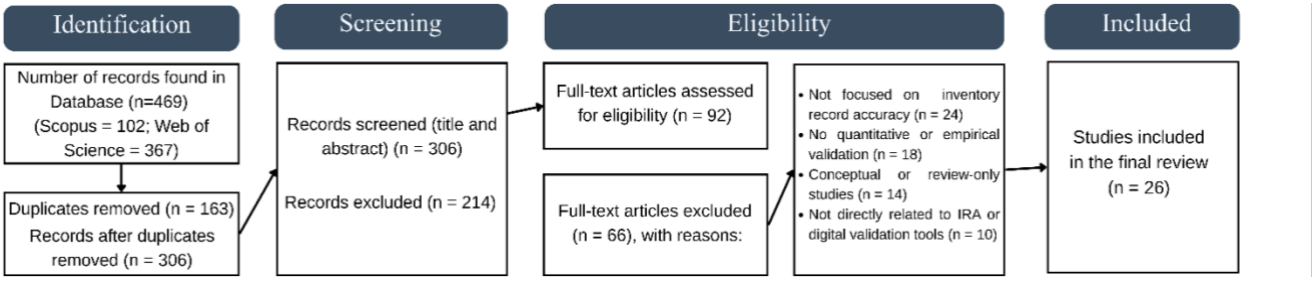


Figure 1: PRISMA, Identification of Studies Through Databases

1.3. Objectives

The main objective of this research is to enhance Inventory Record Accuracy (IRA) in white-goods warehouse operations through the development and validation of an integrated Lean–Digital control architecture that ensures real-time synchronization between physical inventory movements and digital system records.

To achieve this objective, the study systematically analyzes the root causes of inventory discrepancies in current logistics processes, designs a layered solution combining Pick-to-Light, RFID traceability, and digital Poka-Yoke mechanisms, and quantitatively evaluates the operational and economic impact of the proposed model using discrete-event simulation. By measuring improvements in accuracy levels, error reduction, and cost performance, the research aims to assess the effectiveness of integrating Lean principles with digital validation technologies and to provide a replicable framework applicable to similar warehouse environments.

2. Methodology

2.1 Type, design and approach

This study is classified as applied research, as it seeks to design and evaluate a practical solution to improve Inventory Record Accuracy (IRA) in a real warehouse environment through the integration of digital control and traceability technologies [7]. Rather than solely describing existing conditions, the research focuses on improving logistics system performance through a structured intervention. In this context, prior research has demonstrated that inventory inaccuracies are closely associated with process variability, human error, and insufficient control mechanisms within warehouse operations, identifying that these factors are the primary drivers of data corruption in supply chain management [8]. These deficiencies tend to propagate across interconnected activities such as picking, dispatch, and transactional recording, generating cumulative discrepancies that affect overall system reliability and operational performance, especially when manual interventions lack digital verification [9]. Consequently, several studies emphasize the importance of implementing integrated control approaches that combine operational standardization with digital validation technologies to effectively mitigate such errors and stabilize the information flow [10].

From a conceptual perspective, Inventory Record Accuracy (IRA) is defined as the degree of alignment between physical inventory and system records, serving as a critical indicator for warehouse performance

and decision-making processes that determine the financial health of the logistics chain [11]. Achieving high levels of IRA remains particularly challenging in environments characterized by heterogeneous technological adoption and limited process integration, especially in emerging logistics contexts where the gap between operational needs and technological infrastructure is more pronounced [12], [13].

To address these challenges, the literature highlights the role of traceability systems and real-time data capture technologies in improving inventory visibility and control, arguing that real-time monitoring is the only way to eliminate the "black holes" of information in the warehouse [14]. In addition, the implementation of digital solutions has been associated not only with gains in operational efficiency but also with reductions in risk and improvements in the sustainability of logistics operations through the optimization of resources and reduction of waste [15]. From a methodological standpoint, discrete-event simulation is widely recognized as a robust tool for modeling complex warehouse systems and evaluating the impact of process improvements under controlled conditions without disrupting the ongoing physical operation [16]. Furthermore, the integration of enabling technologies such as Pick-to-Light and RFID has demonstrated significant effectiveness in reducing human error, improving process synchronization, and strengthening consistency between physical operations and digital inventory records, validating their role as essential components of a robust "zero-error" logistics strategy [17], [18]. Methodologically, a quasi-experimental design is adopted by comparing two operational scenarios: the current state (As-Is), based on historical data, and an improved state (To-Be) incorporating Pick-to-Light, RFID, and digital Poka-Yoke controls. This comparison enables the analysis of cause–effect relationships between process failures such as picking errors, dispatch overloads, and inconsistent ERP records and the proposed mitigation mechanisms. The study follows a quantitative approach based on the objective measurement of variables, the collection of numerical data, and the statistical analysis of simulation outputs. Key performance indicators include Inventory Record Accuracy (IRA), picking errors, missing units, processing time, and costs associated with rework and returns. Data extracted from ERP logs, discrepancy reports, and audit records were used to parameterize a discrete-event simulation model, allowing controlled experimentation and evaluation of the proposed improvements prior to real-world implementation, ensuring that the model reflects the stochastic nature of warehouse activities and the statistical significance of the results [14], [19]

2.2. Population and sample

The population considered in this study comprises the main operational processes of the white-goods warehouse, including the receipt of finished products, storage and inventory control, order picking, dispatch validation, and administrative registration of inventory movements in the ERP system. These processes represent the complete operational flow through which discrepancies between physical stock and digital records may arise. A non-probabilistic, purposive sampling approach was applied, selecting operational data representative of the warehouse's regular activity. The sample was constructed based on the average monthly flow of approximately 1,200 processed units, with a specific focus on the product line with the highest handling frequency. This approach ensured that the data accurately reflected typical operating conditions. Data collection involved extracting information from ERP records, internal inventory movement reports, picking orders, dispatch validation logs, and discrepancy records. In addition, physical–digital comparisons were conducted to identify differences between executed quantities and system-registered quantities. These datasets provided the basis for quantifying the current level of Inventory Record Accuracy (IRA) and for parameterizing the simulation model used in subsequent stages of the study[19][20].

2.3. Research Variables

The research variables were defined to analyze the causal effect of implementing a Lean–Digital layered control architecture on inventory reliability and warehouse operational performance, the dependent variables

capture changes in inventory accuracy and process stability. Table 1 presents the classification, measurement units, and operational definitions of the independent, dependent variables considered in this study.

Table 1. Research Variables

Variable Type	Variable	Operational Definition
Dependent	Inventory Record Accuracy (IRA)	Percentage of correspondence between physical inventory and ERP system records.
Independent	Picking execution errors	Number of incorrect picking operations detected per operational cycle.
Independent	Picking execution errors	Quantity of items physically unavailable despite system registration.
Independent	Logical–physical warehouse variance	Numerical difference between ERP-recorded inventory (logical stock) and physically verified inventory (physical stock).

2.4 Procedures and methods

The methodological framework was designed to ensure logical traceability between the diagnostic phase, model design, simulation parameterization, and validation of results, enabling a consistent comparison between the current operational condition (As-Is) and the proposed scenario (To-Be) [4]. The study began with a structured diagnosis of warehouse operations, supported by historical operational data, including inventory discrepancies identified during audits, SAP transaction records (MIGO), return documentation, and customer claim reports [4]. A total dataset of 21,507 discrepancy records was analyzed, providing a robust empirical basis for the study. Process flow diagrams were developed to map the operational sequence from physical receipt to final registration in the ERP system, ensuring full visibility of both material and information flows [4].

To support the diagnostic phase, root cause analysis tools were applied. An Ishikawa diagram was used to classify potential causes into categories such as methods, manpower, materials, environment, and information, while Pareto analysis was employed to identify the most significant contributors to inventory inaccuracies [6]. Figure 2 presents the cause–effect relationships derived from the data analysis.

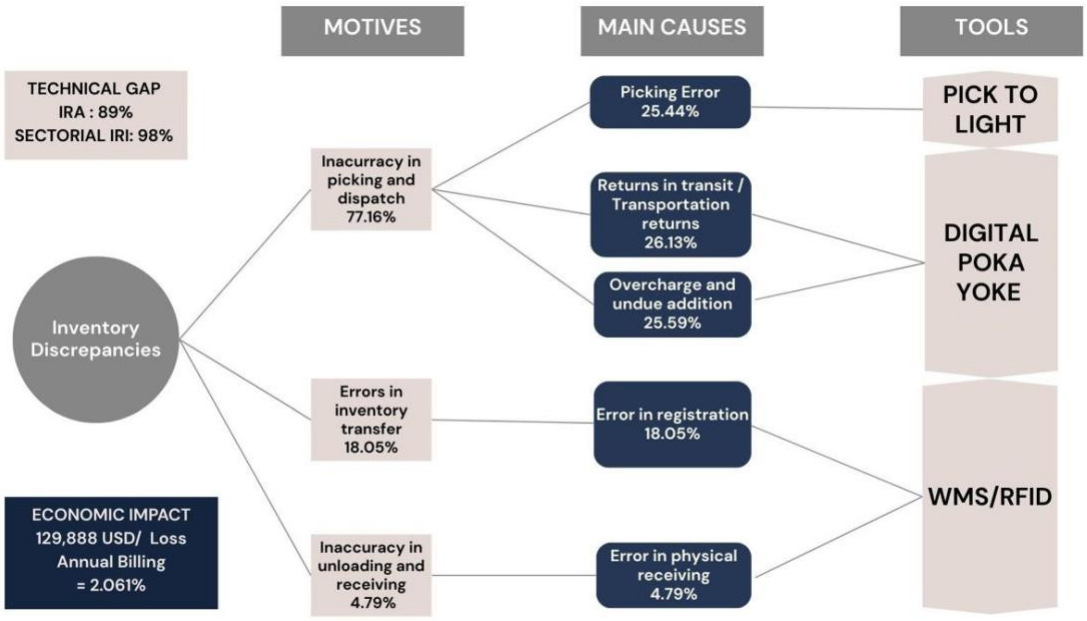


Figure 2: Cause-effect analysis

Based on these findings, an integrated Lean–Digital closed-loop control model was conceptually designed to directly address the dominant sources of variability. The proposed architecture combines Pick-to-Light (PTL) technology to reduce SKU selection errors through visual confirmation, RFID validation with overload alerts to automatically verify dispatch coherence with planned orders [5], and a digital Poka-Yoke mechanism embedded in SAP to block transactional postings when inconsistencies between physical stock and digital records are detected. Unlike conventional warehouse environments that rely primarily on barcode confirmation and subsequent audits, the proposed configuration introduces preventive digital enforcement by ensuring that transactions cannot be completed unless physical and informational flows remain coherent. This approach transforms validation from a corrective mechanism into a real-time control structure, consistent with evidence on automation and Poka-Yoke applications in industrial systems [6].

To evaluate the operational impact of the proposed intervention, a discrete-event simulation model was developed using Arena software. Input parameters were established based on historical ERP logs and structured sampling. From a total population of 21,507 records, a representative sample size of 381 operational units was determined using a 95% confidence level ($Z = 1.96$) and a 5% margin of error. For simulation purposes, this sample was operationalized into a daily workflow configuration of seven containers per cycle, each with a standard capacity of 50 products, representing a processing range of 350–400 units per operational cycle. Processing times were modeled using triangular distributions derived from historical minimum, most likely, and maximum cycle durations. Picking error occurrence was represented using a Bernoulli distribution calibrated to the historical probability (approximately 9%), while RFID reliability was parameterized at a 99% read accuracy level based on industrial benchmarks [3].

Model verification was conducted by validating entity flow logic, event sequencing, and the activation of digital blocking rules under mismatch conditions, ensuring that the simulation accurately replicated both open-loop (As-Is) and closed-loop (To-Be) system behaviors. Model validation was performed by comparing simulated As-Is outputs with historical performance indicators. The simulated IRA (89.4%) closely matched the observed historical value (89%), with a deviation below 1%, confirming an adequate representation of the real system [7]. To assess model robustness, sensitivity analysis was conducted by varying picking error probabilities (±10%) and processing times (±15%). Across all tested configurations,

the To-Be scenario consistently maintained IRA levels above 96%. Each scenario was executed with 100 independent replications to minimize stochastic variability, and the results are reported with 95% confidence intervals to ensure the statistical reliability of the findings.

2.5 Statistical Input Modeling and Simulation Parameterization

To ensure statistical rigor in the simulation model, operational data were analyzed using Arena® Input Analyzer. Goodness-of-fit was evaluated through Chi-square and Kolmogorov–Smirnov tests at a significance level of $\alpha = 0.05$. Distributions with p-values ≥ 0.05 were accepted, as the null hypothesis could not be rejected. The Mean Squared Error (MSE) was used as a complementary selection criterion to support distribution selection. Container arrivals averaged 5.24 units per day. Since the simulation horizon was defined in minutes (600 minutes), the arrival rate was converted into interarrival time to ensure dimensional consistency. The mean interarrival time was calculated as 1440 / 5.24 = 274.8 minutes. Therefore, arrivals were modeled using an exponential distribution EXP(274.8), aligning daily operational data with the minute-based simulation environment. A total of twelve stochastic processes were parameterized in the model, including exponential distributions for arrivals and dispatch operations, normal distributions for unloading and picking activities, uniform distributions for internal transport times, beta distributions for batch variability processes, and lognormal distributions for error-related events (MIGO and picking errors). All selected distributions satisfied the statistical acceptance criteria ($p \geq 0.05$), ensuring that operational variability was adequately represented. The model was executed using 100 independent replications to reduce stochastic variability. Output stability was verified through 95% confidence intervals for the main performance indicators, particularly Inventory Record Accuracy (IRA) and missing units. The As-Is simulation results (IRA $\approx$ 89%) were consistent with historical warehouse data, supporting the external validity of the model.

2.6 Conceptual Model

Based on this validated data structure, the conceptual engineering model (Figure 3) was developed to outline the transformation of the logistics system. It contrasts the current manual configuration (*As-Is*) with the proposed technological integration (*To-Be*). Specifically, the *To-Be* scenario establishes a layered control system driven by the synergistic interaction between Pick-to-Light devices, RFID technology, and a digital Poka-Yoke embedded in the WMS. This multi-level control architecture aligns with the Lean-RFID integration frameworks proposed by [8], which demonstrate that coupling automatic identification technologies with continuous improvement principles significantly enhances both inventory accuracy and overall system efficiency [20] [21].

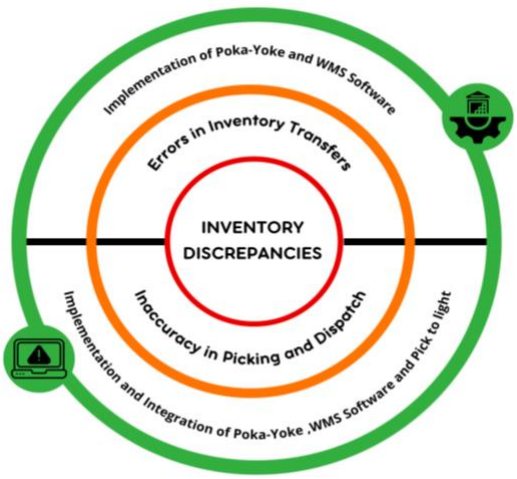


Figure 3: Conceptual Model of the improve scenary

To operate this architecture, the specific logical sequence and control methodology are detailed in the process diagram below. Figure 4 illustrates the complete process flow of the To-Be scenario, providing a comprehensive visualization of the sequential validation steps spanning from the initial order generation to the final system record update. Unlike the traditional open-loop process, this diagram highlights the integration of critical decision nodes where the WMS and RFID technologies actively cross-reference data. This structure ensures data integrity by embedding digital Poka-Yoke constraints at each stage, effectively forcing a real-time verification between physical stock movements and digital records to prevent the propagation of errors before they are committed to the ERP system[18][22].

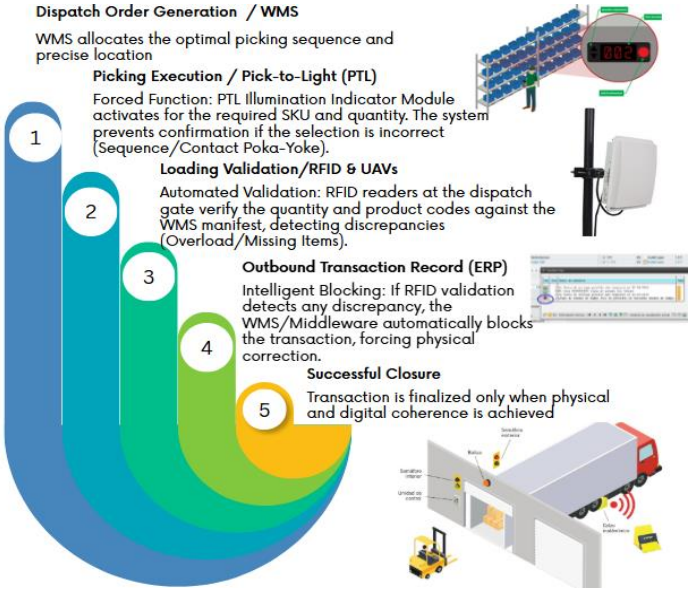


Figure 4: Detailed Process Flow of the Lean-Digital Model

3. Results

3.1. Statistical Distribution Modeling

Operational time data were collected for the main warehouse processes, including container arrivals, unloading, internal transport, picking, dispatch, and transactional discrepancy events. Each dataset was statistically evaluated using Chi-square and Kolmogorov–Smirnov goodness-of-fit tests at a 5% significance level.

For example, the unloading time for stoves was statistically modeled using a continuous probability distribution that achieved a mean squared error (MSE) of 0.004950, indicating acceptable fitting accuracy. Likewise, the transfer time to temporary storage areas was best represented by a Uniform distribution within the interval [12, 43] minutes, with an MSE of 0.006383. The associated goodness-of-fit tests yielded a Chi-square p-value of 0.5 and a Kolmogorov–Smirnov p-value greater than 0.15, as illustrated in Figure 5.

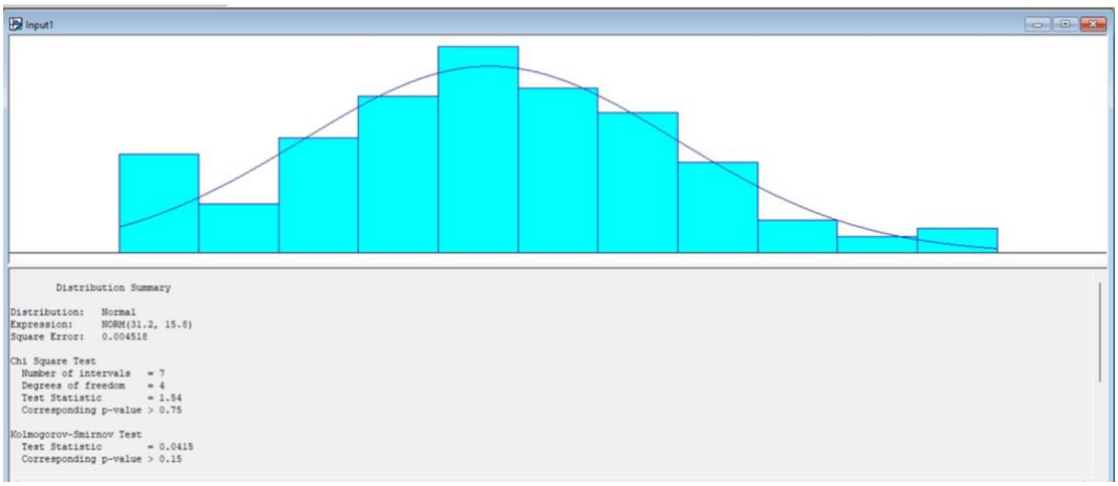


Figure 5: Input Analyzer

All selected distributions presented p-values above 0.05, indicating that the null hypothesis was not rejected and that the probabilistic models adequately capture the observed operational variability. These validated distributions were subsequently embedded into the Arena simulation environment to represent the stochastic behavior of the logistics system.

The initial diagnosis revealed an Inventory Record Accuracy (IRA) of 89%, falling below the 95%–98% industry benchmark. Root-cause analysis of 21,507 incidents traced 77.16% of errors to picking and dispatch operations. The As-Is simulation confirmed significant inefficiencies: 9 (±2) missing units per cycle, an average transfer time of 8.89 h, and resource utilization under 45%, indicating unbalanced workloads, The proposed model integrates Pick-to-Light, RFID for traceability, and a digital Poka-Yoke within SAP (MIGO) to block transactional errors. Validated using Arena simulation software, the optimized operational flow featuring synchronized RFID checkpoints and WMS validation logic is illustrated in Figure 6.

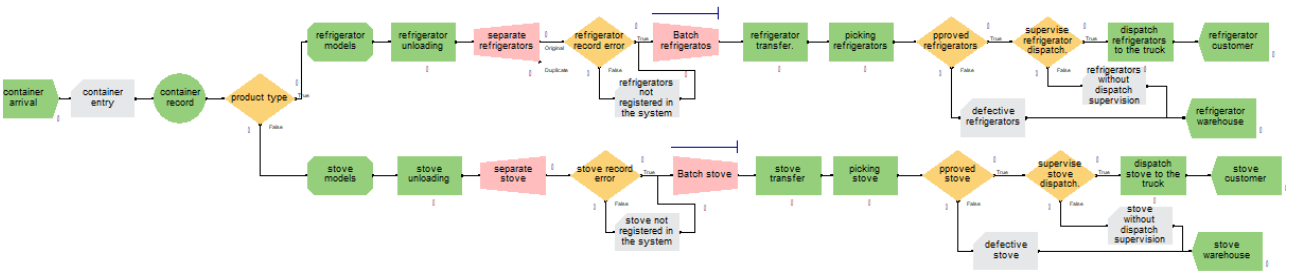


Figure 6: Simulate Model of the improve scenary

Following the implementation of this model, a quantitative assessment was conducted to determine the extent to which the integrated technologies contributed to mitigating the underlying causes of discrepancies. The measurable outcomes derived from this evaluation are systematically presented in Table 2, providing clear evidence of the model's effectiveness in addressing the identified issues.

The simulation model was validated by comparing its outputs with historical operational data, achieving a deviation below 1%, which indicates a high level of representativeness of the real system.

Table 2. Quantitative Impact of the Integrated Model on Inventory Discrepancy Mitigation

Main Causes	Associated Metric	As-Is	To-Be	Relative Impact
Transactional Data Entry Error	SAP System Discrepancies	3 events	2 events	-33.00%
Location and Assignment Error	Errors in Picking Execution	14 events	2 events	-85.70%
Identification and Traceability Failure	Unlocated or Missing Units	9 units	1 unit	-88.90%
Aggregate Inventory Accuracy	Inventory Record Accuracy (IRA)	89.00%	98.30%	9.3 pp

3.2 Economic Impact Estimation

The annual loss of USD 129,888 was calculated based on historical discrepancy records, including the average cost per missing unit, rework labor cost, and dispatch correction expenses. The total annual impact was obtained by multiplying the average number of discrepancy events per year by the estimated cost per event. The projected annual savings exceeding USD 75,000 were derived from the simulated reduction rates observed in the To-Be scenario. Specifically, the percentage reduction in missing units, picking errors, and transactional discrepancies was applied proportionally to the historical cost structure. Therefore, the economic benefit represents an estimated impact based on simulation results rather than a confirmed post-implementation outcome.

4. Discussion

The results indicate that Inventory Record Accuracy (IRA) in retail warehouses is a systemic issue associated with human error, legacy systems, and insufficient prioritization of control mechanisms [8], consistent with studies linking IRA to shrinkage, process variability, and weak control policies [9]. The integration of Pick-to-Light, RFID, and digital Poka-Yoke within a Lean–Digital architecture extends existing research on dynamic reordering and inspection policies under record inaccuracies [9] and RFID-enabled warehouse management [23] by operationalizing IRA through coordinated interventions in picking, dispatch validation, and ERP transaction processes. The resulting IRA level of approximately 98% slightly exceeds international benchmarks for high-performing retail chains[24], reinforcing evidence that levels below 95% increase the risk of stockouts and misaligned replenishment decisions.

However, this study is subject to certain limitations. It is based on one year of data from a single warehouse, which may not capture long-term patterns or rare disruptive events [22], [25]. In addition, the validation relies on discrete-event simulation which, although widely supported in the literature for evaluating warehouse policies [26], may not fully reflect real-world implementation constraints. Future research should extend the proposed model to multi-warehouse environments, incorporate additional performance indicators such as On-Time In-Full (OTIF) [19] and customer experience metrics, and evaluate the integration of advanced Industry 4.0 technologies, including drone-based cycle counting [8] and computer vision systems.

5. Conclusions

This study addresses Inventory Record Accuracy (IRA) as a critical operational and economic challenge in white-goods warehouses, commonly associated with human errors, fragmented internal controls, and weak synchronization between physical inventory movements and digital system records. Through the design and

simulation-based evaluation of a Lean–Digital model integrating Pick-to-Light, RFID traceability, and digital Poka-Yoke mechanisms, the research provides evidence of the potential impact of layered digital controls on inventory accuracy and operational reliability. Simulation results indicate that the proposed model can substantially improve functional performance. In the To-Be scenario, Inventory Record Accuracy (IRA) increased from 89% to approximately 98%, while missing units, picking and dispatch errors, and transactional registration errors (MIGO) were significantly reduced. These improvements suggest that enforcing real-time physical–digital coherence through closed-loop validation mechanisms may effectively mitigate the primary sources of inventory discrepancies identified in the As-Is process.

From a conceptual perspective, this research contributes by formalizing a layered closed-loop control architecture that integrates multiple validation mechanisms into a unified system. Unlike conventional warehouse environments, where controls are applied sequentially or after process execution, the proposed approach enforces real-time transactional consistency, thereby shifting inventory management from error detection to error prevention. From a practical standpoint, the model offers a replicable and adaptable framework for warehouse environments in which manual operations coexist with ERP-based systems and intermediate levels of automation. Nevertheless, the findings are derived from discrete-event simulation and are therefore subject to modeling assumptions, highlighting the need for future research focused on real-world implementation and validation of the proposed architecture.

References

[1] Destro, I. R., Staudt, F. H., Somensi, K., & Taboada, C. (2023). The impacts of inventory record inaccuracy and cycle counting on distribution center performance. Production, 33. https://doi.org/10.1590/0103-6513.20220077

[2] Morenza-Cinos, M., Casamayor-Pujol, V., & Pous, R. (2019). Stock visibility for retail using an RFID robot. International Journal of Physical Distribution & Logistics Management. International Journal of Physical Distribution & Logistics Management (2019) 49 (10): 1020–1042. https://doi.org/10.1108/IJPDLM-03-2018-0151

[3] Kathiriya, A. (2025). A Review of RFID Technology's Impact on Inventory Management and Supply Chain Optimization. International Journal of Emerging Research in Engineering and Technology, Pearl Blue Research Group,Volume 6 Issue 3 PP 86-94. https://doi.org/10.63282/3050-922X.IJERET-V6I3P110

[4] Manuel-Cierto, N., Villayzan-Palomino, J., Saenz-Moron, M., & Straujuma, A. (2025). IoT, RFID, and Lean Tools: An Innovative Strategy to Improve OTIF in the Logistics Chain of a Clothing Marketing Company. 5th LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2025).. https://dx.doi.org/10.18687/LEIRD2025.1.1.306

[5] Wang, Y., & Li, Q. (2025). Optimization techniques in smart logistics: A comprehensive study on machine learning and IoT applications. Proceedings Volume 13792, Fifth International Conference on Intelligent Traffic Systems and Smart City (ITSSC 2025). https://doi.org/10.1117/12.3090136

[6] Acevedo-Aybar, C., Jáuregui-Alfaro, M., Quiroz-Flores, J. C., & Ali, A. (2024). Optimizing warehouse management in footwear commercial companies: a case study on LEAN-BPM. International Journal of Mechanical Engineering by SSRG - IJME Journal, Volume 11 Issue 1. https://doi.org/10.14445/23488360/IJME-V11I1P102

[7] Hofstra, N., Spiliotopoulou, E., & Sander de Leeuw. (2022). Ordering decisions under supply uncertainty and inventory record inaccuracy: An experimental investigation. Decision Sciences, A Journal of the Decision Sciences Institute, Volume 55, Issue 3 Pages 303-318. https://doi.org/10.1111/deci.12564

[8] Fidlerova, H., & Kuka, M., Adamczak, M. (2024). Innovative solutions for warehouse logistics: improving efficiency with RFID and IoT integration. International Scientific Journal about Logistics, Volume: 12 2025 Issue: 1 Pages: 137-145 ISSN 1339-5629. https://doi.org/10.22306/al.v12i1.603

[9] Akkerman, F., Prak, D., & Mes, M. (2025). Dynamic reordering and inspection for the multi-item Inventory Record Inaccuracy problem. European Journal of Operational Research, Volume 321, Issue 2, 1 March 2025, Pages 428-444. https://doi.org/10.1016/j.ejor.2024.09.033

[10] DeHoratius, N., Holpzapfel, A., & co-authors. (2022). Evaluating Count Prioritization Procedures for Improving Inventory Accuracy in Retail Stores. Manufacturing & Service Operations Management, Vol. 25, No. 1. https://doi.org/10.1287/msom.2022.1119

[11] ASTM International. (2017). ASTM E2132-17: Standard practice for evaluation of performance of inventory systems. https://store.astm.org/e2132-11.html

[12] Zhang, A. (2014). Quality improvement through Poka-Yoke: From engineering design to information system design. International Journal of Six Sigma and Competitive Advantage, 8(2), 147-159. https://doi.org/10.1504/IJSSCA.2014.064260

[13] Lazarevic, M., Mandic, J., Sremcev, N., Vukelic, D., & Debevec, M. (2019). A systematic literature review of Poka-Yoke and novel approach to theoretical aspects. Strojniški vestnik – Journal of Mechanical Engineering, 65 (7–8), 443-456. https://doi.org/10.5545/sv-jme.2019.6056

[14] Chehrazi, N. (2024). Exact characterization of the jointly optimal restocking and auditing policy in inventory systems with record inaccuracy. Mathematics of Operations Research 2025 50:1, 656-710. https://doi.org/10.1287/moor.2022.0145

[15] Espinoza Aguirre, J.A., Peña, M., Jadan-Avilés, D., Llivisaca, J. (2022). Relevant Factors in the Inventory Record Inaccuracy for Retail Companies: A Study in Food Retail Industry. Applied Technologies. ICAT 2021. Communications in Computer and Information Science, vol 1535. Springer, Cham. https://doi.org/10.1007/978-3-031-03884-6_40

[16] Castillo, J. (2025). Metamorfosis logística: la evolución de los modelos de distribución para tiendas de proximidad en Latinoamérica. Parte i de la trilogía: transformación de la distribución a tiendas de proximidad en Latinoamérica. Revista Científica Orbis CógnitaVol 9(1):169-184. https://doi.org/10.48204/j.orbis.v9n1.a6718

[17] Łopuszyński, M., Janusz, K., & Karwat, D. (2025). Comparative study of selected order-picking methods: Efficiency, ergonomics, and adaptation rate of new employees. Sensors 2025, 25(3), 923. https://doi.org/10.3390/s25030923

[18] Salehi Shahrabi ,M. (2023). A quantitative analysis of inaccuracy inventory reducing in multi-period mode: Comparison between RFID and inventory counting. Computers & Industrial Engineering, Volume 177, March 2023, 109089. https://doi.org/10.1016/j.cie.2023.109089

[19] Wang, T. (2022). Internet of Things and Big Data in Smart Logistics. Proceedings of the 2022 2nd International Conference on Education, Information Management and Service Science (EIMSS 2022) (pp.1265-1271). https://doi.org/10.2991/978-94-6463-024-4_131

[20] International Organization for Standardization (ISO). (2007). ISO 28000:2007 – Specification for security management systems for the supply chain. https://www.iso.org/standard/44641.html

[21] Vichitkunakorn, P., Emde, S., Masaea, M., Glock, C. H., & Gross, E. H. (2024). Locating charging stations and routing drones for efficient automated stocktaking. European Journal of Operational Research. Volume 316, Issue 3, Pages 1129-1145. https://doi.org/10.1016/j.ejor.2024.03.002

[22] Unhelkar, B, Joshi, S. (2022). Enhancing supply chain performance using RFID technology and decision support systems in the industry 4.0–A systematic literature review. International Journal of Information Management Data Insights, Volume 2, Issue 2, 100084. https://doi.org/10.1016/j.jjimei.2022.100084

[23] Shabani. A., Maroti, G. (2021). Inventory record inaccuracy and store-level performance. International Journal of Production Economics, Volume 235. https://doi.org/10.1016/j.ijpe.2021.108111

[24] Trojanowska, J., Husár, J., & Hrehova, S. (2023).Poka Yoke in Smart Production Systems with Pick-to-Light Implementation to Increase Efficiency. Applied Sciences, 13(17), 11715. https://doi.org/10.3390/app132111715

[25] Barratt, M., Kull, T., Camara, A. (2018). Inventory Record Inaccuracy Dynamics and the Role of Employees within Multi-Channel Distribution Center Inventory Systems. Journal of Operation Management. Volume63, Issue 1, November 2018, Pages 6-24. https://doi.org/10.1016/j.jom.2018.09.003

[26] Vaiciute, K. and Katiniene, A. (2025). Improving the Information Systems of a Warehouse as a Critical Component of Logistics: The Case of Lithuanian Logistics Companies. Systems 2025, 13, 186. https://doi.org/10.3390/systems13030186

Biography

Carlos Ali Irigoin-Sanchez is an Industrial Engineer specializing in Supply Chain Management (SCM) and Global Logistics. His expertise covers Foreign Trade (importation, exportation), local procurement, and strategic negotiation. His research focuses on optimizing warehouse operations and distribution networks using Lean-Digital models specifically integrating RFID and Poka-Yoke principles to enhance data reliability and logistics compliance.

Diego Vincenzo Mascaro-Levano is an Industrial Engineering graduate from the Universidad de Lima, specializing in Supply Chain Management. He has professional experience in logistics and commercial operations within the retail sector, focusing on warehouse management, inventory control, and sales coordination. His primary research interests include Lean Warehousing, digital transformation in logistics, and the application of emerging technologies for inventory optimization.

Jorge Antonio Corzo-Chavez is currently a professor in industrial engineering career at Universidad de Lima and chief executive officer at Conquista Lab. He has experience in advising industrial and service companies to improve productivity and development of business models. His main research areas are in lean manufacturing, master production schedules, logistics, machine learning, and innovation for small and medium-sized enterprises.

Gino Viacava Campos is currently a professor in industrial engineering career at Universidad de Lima. He holds a master's in operations and logistics and is Industrial Engineer from the Universidad de Lima (Peru). Experience as an executive and senior consultant in organizational restructuring situations. His main research interest is lean manufacturing, master production planning, logistics and operations in various sectors.

8th Conference on Production Systems and Logistics

Toward A Coherent Definition Of The Unified Namespace In Industry 4.0

Janis Simons, Fabian Igelmund, Luke Hagenow, Jonas Reck, Niklas Radbruch, Tobias Schröer, Paul Katsioupis

FIR (Institute for Industrial Management) at the RWTH Aachen University, Aachen, Germany

Abstract

The Unified Namespace (UNS) has gained attention in academia and industry as an enabling concept for real-time data integration, event-driven architectures (EDAs), and semantic interoperability in Industry 4.0. However, the concept is marked by divergent interpretations and inconsistent terminology. This paper systematically analyzes existing UNS publications, identifies and classifies recurring characteristics, and proposes a synthesized definition. Establishing a unified definition is essential to provide orientation for the ongoing conceptual discourse. The UNS is defined as a logically centralized, semantically structured integration layer that exposes real-time industrial information across enterprise systems, emphasizing standardized addressing, event publication and subscription, and technology independence to form a consistent basis for research and practice.

Keywords

Unified Namespace; Industrial IT Architecture; Semantic Modelling; Event-Driven Architecture; Industry 4.0

1. Introduction

Industrial production systems are increasingly shaped by automation, digitalization, and heterogeneous cyber-physical resources, resulting in fragmented data landscapes and limited real-time transparency on the shop floor [1], [2], [3]. As manufacturers seek to eliminate data silos, reduce vendor lock-in, and gain control over industrial data, new integration paradigms are emerging to unify shop-floor and IT systems. [3], [4].

The Unified Namespace (UNS) is best understood as an architectural concept that builds upon established integration paradigms such as EDA and semantic data modelling [5], [6]. This concept has gained significant traction in industrial communities, practitioner forums, and vendor ecosystems over the past decade. The term is frequently credited to Industry 4.0 influencer Walker Reynolds and originated within the context of integrating heterogeneous shop floor equipment through publish/subscribe communication mechanisms. [7]

Despite this growing interest and its perceived benefits, there persists a “lack of standard and common definition” [8]. Instead, definitions are highly vendor-specific, community-driven, or embedded in consulting frameworks, leading to inconsistent interpretations across industry and academia. Industry 4.0 requires unified, consistent data structures for real-time control and integration. [9], [10]. The UNS is frequently referenced as a potential enabler of these capabilities, yet it remains insufficiently defined.

Against this backdrop, this paper aims to derive a clear and academically rigorous definition of the UNS. This is done by systematically reviewing existing scientific and practitioner-oriented sources and classifying

DOI: https://doi.org/10.15488/20991
ISSN: 2701-6277

occurring attributes. The classified attributes form the basis of a synthesized and conceptually consistent UNS definition. After this, the UNS is examined in the context of related architectural paradigms, including EDAs, middleware, and the Asset Administration Shell (AAS), highlighting conceptual intersections and complementary roles.

2. Review of Existing Literature

To assess how the concept is currently understood, this chapter distinguishes between peer-reviewed literature, analyzed through a structured review, and practitioner publications, examined via purposeful sampling to reflect its practice-driven origins. This distinction serves to reveal prevailing conceptual tendencies rather than to evaluate either domain normatively.

2.1 The UNS in Academic Discourse

This subsection reviews how the UNS is defined in peer-reviewed academic literature. To ensure methodological transparency, a short systematic literature search was conducted to identify explicit scientific definitions of the term.

The search was performed in the Scopus database using the search string “Unified Namespace” within title, abstract and keywords. Scopus was selected due to its broad interdisciplinary coverage and transparent indexing of peer-reviewed literature. The search string was intentionally kept specific to capture only publications that explicitly focus on the term as a theoretical construct. Results were restricted to English-language publications (n = 27). While three publications could not be retrieved in full text and were therefore excluded from further analysis, subsequent text screening focused exclusively on contributions that provide an attempt at describing the concept of the UNS. In this step, the corpus was reduced to seven publications, three of which were discarded because they were either not peer-reviewed or because the term was only described superficially without sufficient elaboration.

The final sample consists of only four peer-reviewed publications with explicit definitions, reflecting the current state of academic research on the Unified Namespace. For comparability, only structural and conceptual characteristics were considered as definitional elements; benefits, use cases, and implementation details were intentionally excluded.

“A Unified Namespace (UNS) is a data-centric architecture for organizing operational and business data, providing access to all data that clearly and systematically describe the system. This structure makes data available to any device, software, or person who needs it to perform their tasks, using an agreed-upon naming convention and semantic hierarchy.” [11]

Freitas et al. highlight the UNS primarily as a semantically structured information space, emphasizing naming conventions and hierarchical organizations as the core mechanism for data accessibility. This perspective stresses semantic coherence and structural clarity, positioning the UNS as an organized data landscape rather than an integration mechanism or communication paradigm.

“The UNS is an architecture that organizes all entities, resources, and smart devices within a central hub or broker of structured data. This central hub becomes a single source of truth for all organization’s information, providing a comprehensive framework that collects, organizes, and enables real-time access to data at all enterprise levels, facilitating communication between smart devices. Essentially, the UNS represents both the business structure and all its real-time events, acting as the place where the organization’s current state lives. The UNS can translate and contextualize data, facilitating integration across different systems and platforms. [...] The UNS architecture is based on the publish-subscribe communication pattern, so it can be easily integrated using MQTT.” [12]

Díaz et al. conceptualize the UNS as a centralized, broker-based architecture that functions as a single source of truth for the entire organization. Their definition combines structural, semantic, and communicative

dimensions: the UNS is not only a structured data hub that mirrors the business hierarchy and real-time state of the enterprise, but also an integration layer enabling interoperability across heterogeneous systems. They ground the architecture in a publish/subscribe pattern that can be implemented using message protocols such as Message Queuing Telemetry Transport (MQTT).

“The UNS is an event-driven approach to Smart Manufacturing. [...] The resource that produces an event publishes the results to an event broker. Every system interested in a given event can subscribe to the event broker and receive the information related to that event when it occurs. […] The UNS is technology agnostic, since it is a conceptual framework that relies on the event-driven approach. At its core is a naming system inspired by the ISA-95 reference architecture that was discussed previously to ensure that all participants can find the information they are interested in from the event broker.” [13]

Salcher et al. characterize the UNS primarily through event-driven communication principles, highlighting publish/subscribe mechanisms and loose coupling. Their definition positions the UNS closer to EDA, emphasizing reactivity, scalability, and technological independence.

“The UNS consists of a space in the cloud where data can be sent and read in a simple way and based on an open architecture. The UNS concept makes it possible to verify and distribute all data from a single location, so that all elements connected to the UNS are both producers and consumers of existing data in a process. The structure of the UNS proposes the use of only relevant changes, in a way that allows a semantic orientation that responds to events [...].” [14]

Xavier et al. stress the UNS as a shared interaction space in which all entities act as producers and consumers. They only implicitly describe an event-driven, report-by-exception logic. Their perspective underscores dynamic information exchange, selective publication of state changes, and a semantically guided event model.

2.2 The UNS in Industrial Practice

Beyond peer-reviewed contributions, a broad spectrum of industry publications, vendor whitepapers, and practitioner-oriented definitions further shape the conceptual landscape of the UNS. As the term originated in industrial practice before entering academic discourse, these sources are particularly relevant and influence how the concept is currently understood and implemented.

Given the volume and heterogeneity of non-academic sources, a systematic database search was not feasible. Instead, a purposeful sampling strategy was applied: Widely cited and influential practitioner sources were examined. Particular attention was given to a selection of publications that exhibit typical conceptual tensions or ambiguities in relation to the scientific definitions discussed in Section 2.1. These sources are cited throughout the text and included in the attribute-based analysis presented in Table 1. This contrastive selection strategy is not aiming at exhaustiveness but rather highlights areas of divergence and ambiguity across domains. These variations cluster around several dimensions: the conceptual level of the UNS, the granularity of its formal structure, its degree of technological dependence, and its relationship to event-driven architectural principles.

Regarding the conceptual level, some practitioner sources describe the UNS as a concrete infrastructural component. HiveMQ, for example, refers to it as a “central hub or broker for all exchanged information” [15], while Berg characterizes it as a “software layer in the industrial automation system” [5], functionally similar to middleware. In contrast, Díaz et al. intentionally conceptualize UNS as an “architecture that organizes all entities, resources, and smart devices within a central hub or broker” [12], focusing on structural functions rather than specific implementations. Vendor whitepapers such as Capgemini Engineering [8] adopt similar architectural views when discussing UNS-based integration patterns. Across these sources, it remains unclear whether the UNS should primarily be understood as a distinct system component or as a higher-level design principle.

The definitions also diverge in their formal granularity. Harrington, for example, reduces the concept to the "topics and payloads" of MQTT and describes the UNS as "the organization of the payloads (topic hierarchy)" [16]. This interpretation foregrounds syntactic arrangement of messages rather than richer semantic structuring. Freitas et al., by contrast, stress an "agreed-upon naming convention and semantic hierarchy" [11] that systematically organizes operational and business data. Similarly, industry explanations such as i-flow's introductory material [7] and HiveMQ's educational content [17] emphasize consistent structuring of topics and events, yet vary in whether they imply a simple naming scheme or a more explicit semantic model. As a result, the UNS is alternately depicted as a minimal syntactic namespace or as a semantically expressive, context-rich information space.

A further axis of variation concerns technological dependence. Schiffler explicitly associates UNS implementations with MQTT, describing it as "a key protocol for its implementation" [18] and devoting substantial attention to Sparkplug B as a structuring mechanism for UNS topic spaces. Several vendor resources adopt a similar MQTT-centric framing [15], [5]. In contrast, Salcher et al. explicitly claim that "the UNS is technology agnostic" [13]. This tension raises the question of whether specific protocols are intrinsic to the UNS or merely typical realization choices.

Finally, the relation to EDA remains ambiguous. Theocharis explicitly refers to the UNS as "a powerful event-driven architecture" [19], while Xavier et al.'s description of "producers and consumers" in a shared space [14] and Salcher et al.'s event-driven formulation [13] suggest at least an event-driven approach to Smart Manufacturing. Other sources, particularly MQTT-focused introductions [15], [5] describe publish/subscribe mechanisms and real-time updates without consistently embedding the UNS in the broader EDA discourse [20], [21]. As a result, it is not yet clear whether event orientation is a defining characteristic of the UNS or simply a widely adopted implementation pattern.

2.3 Consolidated Results of the Literature Analysis

Existing definitions emphasize important but isolated aspects and lack conceptual integration. Therefore, six recurring attributes are synthesized from the literature to systematically compare interpretations and provide the basis for a consolidated definition. Across the reviewed sources, these attributes are articulated with varying degrees of explicitness: some are stated directly, others are implied through descriptive emphasis, while several attributes are not addressed or cannot be inferred form the respective definitions.

Event orientation (A1) is explicitly emphasized in some of the reviewed scientific sources [13], [14] and is consistently linked to the representation of dynamic system behaviour where the UNS is framed through event-driven principles. In this context, event orientation denotes that information is exposed as events reflecting state changes of assets, processes, or systems rather than as static data records. This framing positions the UNS as a mechanism for propagating operational change in real time, particularly in definitions that build upon publish/subscribe communication and EDAs.

Asynchronous and loosely coupled integration (A2) is frequently described in conjunction with event-oriented approaches, highlighting the decoupled interaction between producers and consumers [13], [18]. In the reviewed literature, this attribute characterizes how systems interact through a namespace layer without direct point-to-point dependencies. While this pattern is emphasized for its contribution to scalability, it is not uniformly treated as a defining property of the UNS itself and appears primarily as an accompanying architectural characteristic of event-based realizations rather than as a necessary conceptual element.

Semantic structuring (A3) refers to the explicit organization of information according to a shared naming convention and contextual hierarchy, enabling consistent interpretation and navigation across heterogenous systems. This attribute differentiates the UNS from generic messaging infrastructures by embedding semantic meaning directly into the namespace structure, as emphasized in definitions that highlight agreed-upon naming conventions and semantic hierarchies. [11]

Closely related to this, **logical centrality (A4)** describes the perception of the UNS as a single, coherent information space that unifies distributed operational and business data. Importantly, this centrality is typically described as logical rather than necessarily physical, meaning that the UNS serves as a unified point of reference even when implementations are technically distributed. [12], [15]

The attribute **technology agnosticism (A5)** addresses the level at which the UNS is conceptualized. While many sources discuss concrete implementations using specific protocols or platforms, the UNS itself is frequently described at an architectural level that remains independent of particular technologies [13]. In this sense, agnosticism denotes that the concept is not inherently bound to MQTT, OPC Unified Architecture, or any other mechanism, even though such technologies are commonly used in practice.

Finally, **unique addressability and subscribability (A6)** emerges as a functional property that enables interaction with the namespace. In the UNS context, this attribute implies that individual information objects, such as asset states, process events, or measurements, can be uniquely identified within the namespace and selectively accessed by authorized consumers. This property is strongly implied by the use of structured namespaces and publish/subscribe mechanisms and underpins the practical usability of the UNS as a shared integration space. [13], [16]

In addition to the six attributes summarized above, the literature occasionally references further characteristics that enrich the conceptual understanding of the UNS without being treated as defining elements. **Cross-level integration capability (A7)** describes the potential of a UNS to span multiple organizational layers, from operational systems to business applications, as implied in definitions that explicitly reference both operational and enterprise data contexts [11], [12]. **State persistence and historical reconstruction (A8)** refers to the retention or reconstruction of system states over time, for example through implementation-level mechanisms such as retained-message patterns in publish/subscribe systems or external storage solutions, especially Historians. [17], [18]

An overview of the identified attributes and their occurrence across the reviewed sources is provided in Table 1. For each source, attributes are classified as *explicitly stated*, *implicitly inferred*, or *not addressed*.

Table 1: Overview of the definition components attributed to the UNS across the reviewed sources

Attribute / Source	Event Orientation (A1)	Asynchronous / Loosely Coupled Integration (A2)	Semantic Structuring (A3)	Logical Centrality (A4)	Technology Agnosticism (A5)	Unique Addressability / Subscribability (A6)
Freitas et al. (2025)	-	-	✓	(✓)	-	(✓)
Díaz et al. (2025)	✓	✓	✓	✓	-	(✓)
Salcher et al. (2024)	✓	✓	(✓)	-	✓	(✓)
Xavier et al. (2024)	✓	(✓)	✓	✓	(✓)	-
HiveMQ (n.y.)	✓	(✓)	(✓)	✓	-	✓
i-flow (2026)	-	-	-	✓	✓	-
Capgemini Engineering (2025)	-	-	✓	✓	✓	(✓)
Schiffler (2025)	-	(✓)	✓	(✓)	-	(✓)
Harrington (2024)	-	-	(✓)	✓	-	✓
Theocharis (2023)	✓	✓	-	-	-	(✓)
Berg (2023)	✓	(✓)	-	(✓)	-	-

✓ = explicitly stated; (✓) = implicitly inferred; - = not addressed

3. Synthesis-Based Proposal for a Definition of the Unified Namespace

Building on the comparative analysis of existing definitions, the chapter first introduces a classification of UNS attributes according to their conceptual relevance. This classification provides the foundation for distinguishing indispensable characteristics from context-dependent architectural properties. Based on this structured consolidation, the chapter subsequently presents a synthesized definition of the UNS that integrates the core aspects identified across the literature while explicitly delineating optional elements.

3.1 Classification of Core and Optional UNS Attributes

Building on the literature review, the conceptual inconsistencies across existing UNS definitions reveal a clear need for structured consolidation. The analysis done in Chapter 2 shows that each definition captures only a partial perspective and leaves open space for interpretation. To consolidate the insights gained from the preceding analysis, the identified attributes of the UNS are classified according to their conceptual necessity and architectural relevance. This classification distinguishes *core* elements, which represent properties that are indispensable for the conceptual integrity of the UNS and must be present in any definition-independent interpretation, from *optional* elements, which describe architectural characteristics that frequently occur in implementations, but whose presence depends on the specific application context or design objectives. Each attribute is linked to the corresponding fragment that will appear in the synthesized definition. Table 2 provides an integrated overview of this mapping.

Table 2: Classification of UNS Attributes and their Incorporation into the Synthesized Definition

Attribute	**Classification**	**Definition Snippet**
A1 Event Orientation	Core	"Systems publish state changes as events and any authorized application may consume them in real time."
A2 Asynchronous / Loosely Coupled Integration	Optional	"In practical implementations, the UNS is often supported by asynchronous and loosely coupled communication patterns."
A3 Semantic Structuring	Core	"…a shared, semantically structured and contextualized namespace…"
A4 Logical Centrality	Core	"…provides a logically centralized organization of information…"
A5 Technology Agnosticism	Core	"…while remaining independent of specific communication protocols or vendor technologies."
A6 Unique Addressability / Subscribability	Core	"…that enables standardized addressing, event publication, and event subscription through topics..."
A7 Cross-Level Integration Capability	Optional	"…may span both OT and IT domains..."
A8 State Persistence / Historical Reconstruction	Optional	"…may incorporate state persistence or historical reconstruction."

3.2 Synthesized Definition of the Unified Namespace

Based on the classified attributes the following definition is proposed:

A Unified Namespace (UNS) is an enterprise-wide conceptual architectural principle for exposing real-time industrial information within a shared, semantically structured, and contextualized namespace. It provides a logically centralized organization of information that enables standardized addressing, event publication, and event subscription trough topics, while remaining independent of specific communication protocols or vendor technologies. Systems publish state changes as events and any authorized application may consume them in real time.

In practical implementations, the UNS is often supported by asynchronous and loosely coupled communication patterns. Furthermore, it may span both OT and IT domains and may incorporate state

persistence or historical reconstruction. These characteristics are commonly observed in implementations and enhance applicability, but they are not considered constitutive elements of the conceptual definition.

4. Architectural Relationship to Adjacent Concepts and Technologies

This chapter examines the architectural relationship between the UNS and EDAs, middleware technologies, and the AAS, highlighting their respective roles across different architectural layers in order to avoid conceptual overlap and misinterpretation.

It is important to note that the UNS is not intended as an alternative architectural paradigm to these concepts. Rather, it represents an architectural organization and exposure pattern that is typically realized through middleware technologies, operationalized using event-driven interaction principles, and may incorporate standardized semantic information models such as the AAS.

4.1 Relationship to Event-Driven Architecture

EDA is an architectural concept in which systems automatically react to certain events and communicate with each other in a loosely coupled manner. An event could, for example, be "a state change of a component or the attainment of information by a service" [20]. It consists of topics and messages with a payload with the necessary data as well as the context it has occurred in. [20] Systems can publish any relevant event to a specific topic in a broker, and any other system interested can subscribe to this topic to always receive the data in real time. Not only is this more efficient, but since the systems only need to interact with the broker, they do not even have to "know" each other. [22] This asynchronous way of communication makes data integration and upscaling considerably easier [23].

The UNS builds on EDA communication principles and enforces a standardized namespace for organizing event streams. This namespace provides a unified semantic structure that ensures consistent naming, contextualization, and categorization of events. By defining where each data point belongs within the organizational information space, the UNS enables a single, contextualized point of access to distributed data sources and thus complements event-driven mechanisms.

4.2 Relationship to Middleware

Middleware is a software layer that enables communication and coordination between independent systems by managing the transport, routing, and delivery of data, often through message-oriented technologies such as Kafka or RabbitMQ. Related concepts, such as Enterprise Application Integration, pursue similar goals by reducing redundancy, ensuring consistency, and linking heterogeneous systems via intermediary layers. [21], [24], [25], [26] Within EDAs, middleware facilitates loosely coupled interaction by providing a shared exchange layer for event transport across heterogeneous technologies [27].

While the UNS may appear similar to middleware, it operates at a different architectural layer by introducing a logically centralized and standardized semantic structure for organizing industrial information. Middleware typically enables technical interoperability and message transport, whereas the UNS establishes a shared conceptual data model that ensures consistent interpretation of information across the enterprise. [28]

4.3 Relationship to Asset Administration Shells

According to Mäule et al., the AAS is a standardized digital representation of a physical asset that structures asset information into ontology-linked semantic submodels, ensuring consistent interpretation across systems [29]. It provides a uniform, machine-readable interface through which assets expose properties, events, and capabilities.

Mäule et al. show that the AAS aligns well with the UNS, as both aim to structure and standardize information flows in event-driven production environments. Mapping AAS submodels into UNS

publish/subscribe structures enables semantically enriched, hierarchically organized, and real-time accessible asset data [29] complementing the semantic orientation of the UNS described by David et al. [28].

However, the concepts differ in scope. While the UNS provides a centralized, hierarchical data space without prescribing asset-level semantic models [28], the AAS enforces standardized, ontology-based asset representation [29]. Thus, AAS content can populate a UNS, but the UNS does not replace the asset-level semantic rigor of the AAS.

4.4 Summary and Conceptual Positioning

The conceptual distinctions discussed in this chapter substantiate the synthesized definition of the UNS introduced in Chapter 3. They position the UNS as an enterprise-wide conceptual architectural principle that establishes a logically centralized organization of data objects within a shared, semantically structured namespace. This clarifies that the UNS is not defined by specific technologies or communication mechanisms, but by the logical and semantic principles governing how industrial information is organized and accessed in real time.

While EDA defines interaction mechanisms, middleware provides technical connectivity, and the AAS specifies asset-level semantic representations, the UNS operates as an overarching organizational pattern that structures and exposes such information within a logically unified enterprise-wide namespace.

5. Conclusion and Outlook

This paper set out to clarify the conceptual foundations of the Unified Namespace, a term that has seen rising prominence in industrial practice but remains inconsistently defined in academic and practitioner literature. By proposing a definition, future research should focus on several key areas that are essential for advancing both the theoretical understanding and practical application of the UNS. These include empirical validation of the proposed definition through industrial case studies and cross sector comparisons, as well as further development of semantic standardization through shared ontologies and interoperability mechanisms.

Another key focus is the integration of planning and control logic, particularly production planning and scheduling, which is critical because it translates real-time events into coordinated operational decisions affecting schedules, resources, and delivery performance, into event-driven data models with clearly defined event semantics and granularity. In addition, further work is needed to clarify the positioning of the UNS within industrial data architectures, including its relationship to Data Spaces, Data Lakes, Data Warehouses, and Digital Twins, as well as to address open questions related to lifecycle management, governance, versioning, ownership, quality assurance, and the long-term evolution of namespace structures.

Finally, future research should explicitly ground the UNS in concrete industrial use cases (e.g., OEE calculation, predictive maintenance), systematically specifying the required namespace structures, semantic elements, and complementary building blocks or systems, as such use-case–driven analyses can substantially lower adoption barriers for companies and enable the generation of clearly measurable business value.

Acknowledgements

This work was conducted within the framework of the research project nexAMo, funded by the Federal Ministry for Economic Affairs and Climate Protection (Bundesministerium für Wirtschaft und Klimaschutz, BMWK).

References

[1] Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K., 2016. Cyber-physical systems in manufacturing. CIRP Annals 65 (2), 621–641.

[2] Kagermann, H., Wahlster, W., Helbig, J., 2013. Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0.

[3] Leitão, P., Colombo, A.W., Karnouskos, S., 2016. Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges. Computers in Industry 81, 11–25.

[4] Kagermann, H., Anderl, R., Gausemeier, J., Schuh, G., Wahlster, W., 2016. Industrie 4.0 im globalen Kontext – Strategien der Zusammenarbeit mit internationalen Partnern (acatech STUDIE). Herbert Utz Verlag, München.

[5] Berg, C., 2023. Unified Namespace: What is UNS? (Examples and Tips). https://www.clarify.io/learn/unified-namespace-uns. Accessed 25 November 2025.

[6] Pintilie, A., Poelarends, R., Capiluppi, A., 2025. Optimising IIoT Control Systems at Demcon: Integrating MQTT, Sparkplug B, and ISA-88 for Unified Automation, in: Proceedings of the 20th International Conference on Evaluation of Novel Approaches to Software Engineering, Porto, Portugal. April 4-6. SCITEPRESS - Science and Technology Publications, pp. 351–358.

[7] i-flow, 2025. Was ist der Unified Namespace (UNS)? Einfach erklärt. https://i-flow.io/ressources/was-ist-der-unified-namespace-uns-einfach-erklaert/. Accessed 3 December 2025.

[8] Capgemini Engineering, 2025. Unified Namespace - implementation using OPC UA and MQTT. https://www.capgemini.com/fr-fr/wp-content/uploads/sites/6/2025/03/UNS-implementation-using-OPC-UA-MQTT.pdf. Accessed 6 December 2025.

[9] Zuehlke, D., 2010. SmartFactory—Towards a factory-of-things. Annual Reviews in Control 34 (1), 129–138.

[10] Koren, I., Jarke, M., Michael, J., Heithoff, M., Tacke Genannt Unterberg, L., Stachon, M., Rumpe, B., van der Aalst, W.M.P., 2024. Navigating the Data Model Divide in Smart Manufacturing: An Empirical Investigation for Enhanced AI Integration, in: van der Aa, H., Bork, D., Schmidt, R., Sturm, A. (Eds.), Enterprise, Business-Process and Information Systems Modeling, vol. 511. Springer Nature Switzerland, Cham, pp. 275–290.

[11] Freitas, L., Silva, M., Vale, G., Avram, C., Lopes, H., Pereira, F., Leal, N., Machado, J., 2025. OPC UA and MQTT performance analysis within a unified namespace context. Internet of Things 33, 101734.

[12] Díaz, J.A., Rey, J.M., Hernández, I., Duarte, N., Reynolds, W., 2025. Unified Namespace: A New Paradigm for Industrial IoT Data Integration. IEEE Internet Things Mag., 1–10.

[13] Salcher, F., Finck, S., Hellwig, M., 2024. A Smart Shop Floor Information System Architecture Based on the Unified Namespace, in: 2024 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC), Funchal, Portugal. June 24-28. IEEE, pp. 1–9.

[14] Xavier, R.N.G., Cuarelli, C., Kaneshiro, P.J.I., Asato, O.L., Moro, J.R., Nakamoto, F.Y., 2023. Architecture Proposal for SMT Production Line in the Context of Industry 4.0, in: 2023 15th IEEE International Conference on Industry Applications (INDUSCON), São Bernardo do Campo, Brazil. November 22-24. IEEE, pp. 518–524.

[15] HiveMQ, n. y. Unified Namespace (UNS) Essentials. https://www.hivemq.com/mqtt/unified-namespace-uns-essentials. Accessed 25 November 2025.

[16] Harrington, J., 2024. UNS 101: Understanding the Unified Namespace. https://www.iiot-world.com/smart-manufacturing/uns-101-understanding-the-unified-namespace/. Accessed 25 November 2025.

[17] HiveMQ, 2025. A Beginner's Guide to Unified Namespace for IIoT | IoT For All. https://www.iotforall.com/a-beginners-guide-to-unified-namespace-uns-for-iiot. Accessed 3 December 2025.

[18] Schiffler, A., 2025. Implementing Unified Namespace (UNS) using MQTT. https://cedalo.com/blog/uns-mqtt-implementation/. Accessed 25 November 2025.

[19] Theocharis, J., 2023. 3.1.1 Unified Namespace. https://learn.umh.app/lesson/introduction-into-it-ot-unified-namespace/. Accessed 25 November 2025.

[20] Kremer, S., Mennerich, C., Stroh, M.-F., Schuh, G., 2023. Proposing a Solution for a Self-Managed Data-Ecosystem in Production: Use-Case-Driven IT-OT-Integration with an Event-Driven IT-Architecture, in: Proceedings of the Conference on Production Systems and Logistics: CPSL 2023 - 2, Stellenbosch, South Africa. November 14-17. Hannover : publish-Ing.

[21] Thondalapally, K.R., 2025. Event-Driven Architectures: The Foundation of Modern Distributed Systems. IJSAT 16 (1).

[22] Kremer, S., 2023. Der Game-Changer für die Produktion der Zukunft / The Game Changer for the Production of Tomorrow. UdZ – The Data-driven Enterprise 2 (2), 14–19.

[23] Igelmund, F., Tücks, G., 2025. Event-Driven Architecture: Der Turbo für die smarte Fabrik. Factory Innovation, 60–66.

[24] Kodakandla, N., 2022. How Middleware Powers Real-Time Data in Today's Event-Driven World. TIJER - TIJER - INTERNATIONAL RESEARCH JOURNAL (www.TIJER.org), ISSN:2349-9249 9 (8), 1–7.

[25] Lehner, F., Wildner, S., Scholz, M., 2008. Wirtschaftsinformatik: Eine Einführung, 2. Aufl. ed. Hanser, München, 417 pp.

[26] Ferstl, O.K., Sinz, E.J., 2013. Grundlagen der Wirtschaftsinformatik, 7., aktualisierte Aufl. ed. OLDENBOURG WISSENSCHAFTSVERLAG, Berlin, München, 543 pp.

[27] Pramukantoro, E.S., Anwari, H., 2018. An Event-based Middleware for Syntactical Interoperability in Internet of Things. IJECE 8 (5), 3784–3792.

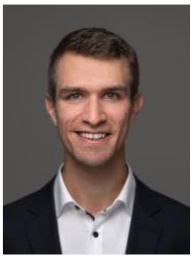
[28] David, J., Martikkala, A., Lobov, A., Lanz, M., 2019. A unified ontology namespace for enterprise integration - a digital twin case study, in: The V International Forum "Instrumentation Engineering, Electronics and Telecommunications - 2019", Izhevsk, Russia. November 20–22. Publishing House of Kalashnikov ISTU, pp. 13–22.

[29] Mäule, J., Kutzler, T., Leich, T., Schäfer, D., Richter, C., 2024. Unified Namespace and Asset Administration Shell: A Winning Combination for Digital Production, in: 2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA), Padova, Italy. September 10-13. IEEE, pp. 1–7.

Biography

Janis Simons, M.Sc. (*1994) is Team Lead within Production Management and PhD candidate at the FIR e.V. at RWTH Aachen University since 2022. As a senior project manager, Simons conducts research and industry projects in the field of business applications like ERP, APS and MES, Unified Namespace concepts, and IT-OT integration.

Fabian Igelmund, M.A. (*1995) is Team Lead of Information Technology Management and PhD candidate at the FIR e.V. at RWTH Aachen University since 2024. As a senior project manager, Igelmund conducts research and industry projects in the field of Event-Driven Architectures, Unified Namespace concepts, and IT-OT integration for industrial environments.



Luke Hagenow, M.Eng. (*2002) is part of the team Production Management and PhD candidate at the FIR e.V. at RWTH Aachen University since 2025. As a project manager, Hagenow conducts research and industry projects in the field of Industrial process optimization, Unified Namespace concepts, and IT-OT integration for industrial environments.

Jonas Reck, M.Sc. (*1992) is Senior Project Manager as part of the group Production Management at the FIR e.V. at RWTH Aachen University since 2023. As a senior project manager, Reck conducts research and industry projects in the field of business applications like ERP and MES, with the focus on EAM, integration patterns, middleware concepts and data ontologies.

Niklas Radbruch, M.Sc. (*1989) is Senior Project Manager as part of the group Production Management at the FIR e.V. at RWTH Aachen University since 2024. In his role he leads research and industry projects focused on business applications like PLM, ERP and MES, with a particular emphasis on EAM, data ontologies and product centric architectures.

Tobias Schröer, M.Sc. (*1992) is Deputy Managing Director of the FIR e.V. at RWTH Aachen University and has served as Head of the Production Management department since 2020. In this role, he is responsible for the strategic and operational development of research and industry projects focusing on production planning and control, business application systems such as ERP, APS, and MES, as well as data-driven and integrated production management concepts.

Paul Katsioupis, B.Sc. (*1998) is student of Electrical Power Engineering and Business Administration. In his role as a Research Assistant in Production Management at the FIR e.V. at RWTH Aachen University, he is occupying himself specifically with the concept of the Unified Namespace in industrial IT infrastructures.

8th Conference on Production Systems and Logistics

Cross-Country Analysis of Digitalization of Services of General Interest: A Systematic Literature Review

Eva Paulina Ziehm, Nick Große, Jonas Eichholz, Leon Jungh, Marius Klinsmer Tchapga Leumen

Chair of Enterprise Logistics, TU Dortmund University, Germany

Abstract

Services of general interest (SGI) act as enablers for ecosystems that provide industrial services, maintain essential supply structures, and support social stability. Yet they face pressure from digitalization, globalization, and demographic changes, with digitalization playing a critical role in improving efficiency. In this context, technology-driven stakeholders may engage with SGI markets, potentially influencing implementation and fostering innovation. Despite these opportunities, SGI companies have difficulty effectively leveraging digitalization, as best practices are difficult to generalize globally due to cultural, demographic, and political differences, revealing a research gap in the cross-country adaptability of SGI digitalization.

This paper addresses this gap by systematically analyzing the structural and institutional factors that affect the implementation and efficiency of digitalization initiatives. Later, these are evaluated to derive implications for the comparability of diverse digitalization approaches, within which digital technologies, including emerging applications such as Artificial Intelligence (AI), constitute integral components, although their implementation in SGI remains underrepresented in the literature. Against that backdrop, this paper addresses the central research question “Which factors influence the implementation of SGI digitalization, and how do these factors affect the transferability of different digitalization approaches from a cross-country perspective?”.

The authors use secondary data from a systematic literature review. They analyze and compare SGI areas using defined criteria, such as organizational form and digitalization level. The study shows that digitalization depends on institutional openness and investment models. Differences appear in e-government maturity, partnership efficiency, and overall system performance.

International comparisons highlight the influence of governance structure on the performance, efficiency, and innovation of SGI. The results contribute to a more nuanced understanding of SGI digitalization by identifying transferable patterns and structural constraints, and they support policymakers and technology providers in designing resilient, context-sensitive digital transformation.

Keywords

Services of General Interest; Cross-Country Analysis; Digitalization; Literature Review; International Comparison

DOI: https://doi.org/10.15488/20992
ISSN: 2701-6277

1. Introduction

Services of general interest (SGI) constitute the essential infrastructure of modern societies, sustain social organization, and ensure the continuous provision of critical supply networks that uphold social stability [1]. SGI play a central role in granting access to vital services, fostering regional development, and maintaining the reliable functioning of public infrastructure [1]. However, SGI faces growing challenges arising from technological, demographic, and global transformations, as well as increasing competition from value-driven industrial enterprises [2].

Digitalization has emerged as a critical driver for improving operational efficiency, service quality, and data-driven decision-making [3]. Despite the recognized potential of digital transformation, the extent of digitalization within SGI varies considerably between regions and countries [4]. Ethical and institutional factors heavily influence the success of such initiatives [5]. In contrast to industrial companies, where faster innovation cycles and higher Artificial Intelligence (AI) adoption rates are evident, AI remains underrepresented in SGI [6].

This imbalance unveils a significant research gap in understanding how structural and institutional conditions shape digitalization outcomes across national contexts. Like small and medium-sized enterprises in the private sector, SGI often lack the capacity to implement digital solutions due to limited financial and human resources, even though such technologies offer clear advantages for monitoring operations, enhancing service quality, and personalizing user experiences [7]. Moreover, municipal SGI are particularly affected by demographic change and global market pressures, making digital integration a matter of both efficiency and long-term resilience.

In this context, the paper explores the research question: "*Which factors influence the implementation of SGI digitalization, and how do these factors affect the transferability of different digitalization approaches from a cross-country perspective?*" The study aims to assess and compare existing levels of digitalization, examining how variations in governance, investment strategies, and institutional openness affect performance and innovation outcomes. To answer the question, our research elaborates on substantial cross-national differences. First, it develops a structured analytical framework of key factors influencing SGI digitalization, including e-government development, partnership efficiency, and overall system performance. Second, it introduces a comparative logic that combines a European and an international perspective, enabling the systematic identification of similarities and differences across institutional contexts. Third, it derives implications for the transferability of digitalization approaches, emphasizing the influence of institutional and structural conditions on digitalization success. The paper is structured as follows: Section 2 provides an overview of the definition, role, and structural characteristics of SGI, highlighting their digitalization context, and explains institutional conditions. Subsequently, Section 3 introduces the systematic literature review used to extract the information base, whose findings will be discussed in Section 4. Section 5 summarizes essential key findings of the research and reveals open issues for further research.

2. Related background

The European Union defines SGI as services that meet general societal needs and contribute to social and economic cohesion [1]. Its omnipresence across various characteristics allows SGI to be differentiated based on factors such as its economic intentions, accessibility, provision area, contribution to territorial capital, effective range, regulatory limitations, its impact on demographic sensitivity, and primary objective [8]. However, there is no uniform international definition of SGI. According to the European Union, they contain essential services indispensable for modern societies, including energy and water supply, transport and mobility, health and education services, as well as communication and digital infrastructures [1]. Divergent responsibilities and institutional structures significantly shape performance, quality, and innovative capacity across SGI sectors [4,9].

In the context of SGI, digitalization refers to the improvement of public services and infrastructures. This includes the digitalization of administrative processes, data-sharing systems, digital citizen services, and increasingly AI-driven applications [4,10]. The aim is to increase efficiency, improve sustainability, and ensure resilience [11]. Compared to the private sector, digitalization in SGI sectors often progresses more slowly, due to strict regulatory frameworks, fragmented responsibilities, and limited financial and human resources [4]. The pace and quality of digital development in SGI are shaped by institutional and structural conditions, leading to considerable differences in the degree of SGI digitalization [4,11]. *Governance structures* determine responsibilities for coordination, performance, and service delivery [12]. Similarly, *funding mechanisms and financial instruments* have an impact on the digitalization capacity of SGI providers [11]. Moreover, private-sector actors play an important role, particularly in providing technological services and Public Private Partnerships (PPP) [4,13]. Private-sector actors refer to market-oriented companies and organizations that provide technical expertise, innovation capacity, and infrastructure beyond the capabilities of public authorities [13]. *Cooperation with private service providers*, like Public Private Partnerships, is understood as long-term cooperative arrangements between public bodies and private firms, characterized by shared responsibilities, risk distribution, and contractual performance agreements in the provision of public services or infrastructure [13].

3. Methodology

Conducting both a European (EU) comparison and a broader international comparison is essential, as EU member states share common regulatory frameworks and policy instruments that shape SGI digitalization, while the international perspective reveals alternative governance models, innovation pathways, and structural conditions beyond the EU context. Together, these complementary comparisons enable a more precise identification of institutional drivers and barriers and strengthen the analytical validity of the transferability assessment.

To investigate the structural and institutional factors affecting the digitalization of SGI and the transferability of digital approaches across national contexts, this study conducts a systematic literature review (SLR). The review follows the methodological framework proposed by vom Brocke et al. [14] to ensure transparency, reproducibility, and a rigorous synthesis of existing knowledge. The scope and objectives of the review are specified using Cooper's taxonomy [15], adopting a neutral perspective, considering a representative, non-exhaustive excerpt of research data, and shedding light on the applicability of results for scholars and practitioners. The SLR combines two complementary perspectives: First, a European comparison of municipal SGI, focusing on countries that operate within shared EU policy frameworks and regulatory standards; Second, an international comparison of SGI, expanding the analysis to non-European contexts to capture alternative governance models, innovation dynamics, and structural conditions. Superimposing both perspectives contribute to both i) strengthening the analytical validity of cross-national comparison and ii) identifying transferability of potentials for digital solutions for SGI.

A multistage search process was conducted to identify scientific publications addressing SGI characteristics, governance structures, and digital transformation outcomes. In an initial exploratory phase, the research topic was conceptually delineated. Two targeted search strings were developed and applied in the Scopus database, selected for its comprehensive coverage of peer-reviewed literature and flexible search operator functions that are shown in Table 1. The first operators contain the main topic of the search, including the most important synonyms of SGI, whereas the second set focuses on cross-country comparisons. The final operator narrows the scope by limiting the search to the EU level or enabling an international comparison.

Table 1: Perspectives, composition of search strings, and literature selection

Perspective	Search string composition	# Paper	Selection (Abstract)	Selection (Fullpaper)
European comparison	TITLE-ABS-KEY(("service* of general interest*" OR "public service*" OR "municipal service*") AND ("compar*" OR "analys*" OR "benchmark*" OR "evaluation") AND ("europ*" AND ("countr*" OR "nationwide")))	799	616	28
International comparison	TITLE-ABS-KEY(("service* of general interest*" OR "public service*" OR "municipal service*") AND ("compar*" OR "analys*" OR "benchmark*" OR "evaluation") AND ("international" OR "nationwide"))	1515	86	47

Both searches were limited to publications written in English. Articles without abstracts, inaccessible sources, and duplicates were excluded to ensure consistent and comparable academic evidence. Titles and abstracts of the remaining publications were screened for relevance to the overarching research question, focusing on studies that analyze SGIs comparatively and address digitalization or structural determinants of SGI. The European search yields an initial set of 799 publications, reduced to 28 relevant sources after language filtering, screening, and backward search additions. The international search covered 1,515 results, from which 47 sources remained after the full selection process. The international search identified comparative studies extending to both EU and non-EU countries. To prevent duplication, all EU-related findings from the international review were systematically included in the European comparison. Furthermore, sources offering only limited information on a single influencing factor were excluded, as they did not allow comparison. Applying these criteria resulted in the inclusion of 14 sources in the international comparison and 16 sources in the European comparison. Both reviews followed the same inclusion criteria, ensuring comparability of findings across national perspectives. Additional grey literature sources, such as reports from international organizations (e.g., OECD, United Nations), were used complementarily to provide contextual data, benchmarking indicators, and insights on the latest global digitalization trends.

4. Findings and Discussion

As part of the systematic literature review, the fine selection serves as an entry point for synthesizing the relevant body of knowledge on the structural and institutional factors of SGI digitalization across countries as well as for discussing the empirical results of the comparative assessment of SGI digitalization. Analyzing and visualizing the findings through a concept matrix based on Webster and Watson (2002) [16], allowing a systematic classification based on the influence factors of SGI digitalization introduced in Section 2, forming the basis for the cross-country comparison. Each publication was analyzed based on content addressing the aforementioned influencing factors on digitalization: i) *degree of digitalization*, ii) *governance structure*, iii*) funding mechanisms and financial instruments*, iv) *size and organizational structure of municipalities*, and v) *cooperation with private-sector actors*. The concept matrices visualized in Table 2 and Table 3 highlight these structural determinants and allow assessing the transferability of different digitalization approaches. A concluding section synthesizes these comparative findings to identify overarching patterns that explain how the influencing factors shape successful digital transformation in SGI.

4.1 European Comparison

The analysis for the European comparison, illustrated in a concept matrix in Table 2 and visualized in an extended version in Table 4 in the appendix, reveals substantial variation in digital development across Europe. Northern European countries such as Denmark, Estonia, Finland, and Sweden have a high degree of digitalization, characterized by advanced e-government infrastructures, strong interoperability, and general acceptance of digital services [4,11,17]. As reflected in the United Nations E-Government Development Index (EGDI), which measures e-government maturity based on online service provision, telecommunication infrastructure, and human capital, Denmark and other Nordic countries are identified as

global leaders in e-government, achieving scores up to 0.99 out of 1 [11]. In contrast, Central European states such as Germany and the Netherlands exhibit slower digital progress, due to fragmented administrative structures and persistent concerns about privacy and user acceptance [18]. The Czech Republic also shows a comparatively limited degree of digitalization, driven by capability constraints, including shortages of specialized personnel and restricted local innovation capacity [19]. These disparities are shaped by governance structures (ii).

In most European countries, local authorities are responsible for key public services, including education, health and social services. High levels of local autonomy, embedded in well-coordinated national frameworks, for example, in Northern Europe, enable individual adaptation and fast innovation cycles promoting digitalization [18,20-23]. Moreover, countries with greater fiscal autonomy at the local level, such as Germany, which allocates more than 10 % of GDP to local public expenditure, or Sweden, where local public expenditure accounts for approximately 24 % of GDP, can sustain continuous investment in digital modernization [20,23]. In contrast, centralization without adequate local autonomy, as observed in parts of Eastern Europe, such as the Czech Republic, tends to result in slower implementation and weaker interoperability among e-government systems [24]. Financial resources (iii) further play a decisive role, as higher-income countries generally achieve more advanced levels of digitalization. In addition, higher levels of municipal spending capacity and financial independence, when combined with local autonomy (ii), significantly enhance digitalization outcomes, as illustrated by the case of Denmark [11,20,25].

Municipal structures (iv) in Northern European countries such Denmark and Finland are highly social-oriented and citizen-focused [20]. Their local governments prioritize sustainability and social policies with high-quality public services [20]. At the same time, municipalities like Sweden retain strong local autonomy while cooperating through inter-municipal services to deliver efficient and equitable solutions [22].

Table 2: Comparison of European Countries

Country	Degree of digitalization (i)	Governance structure (ii)	Financial resources (iii)	Municipal structure (iv)	Private Co-operation (v)
Denmark	- Highest digitalized e-government (EGDI: 0.99/1) [11] - Digital portals for citizen & business, AI implementation [4]	- Decentralized and Local autonomy [20]	- Financial independent [20] - High-income country [11]	- Citizen-oriented, sustainable, and social policy [20]	- Direct award for PPP contracts [26]
Estonia	- High digitalized e-government (EGDI: 0.97/1) [11] - Online services, AI implementation [4]	- Local autonomy [20]	- High-income country [11]	- Citizen included in decision-making [20]	- Direct award for PPP contracts [26]
Finland	- High digitalized e-government (EGDI: 0.96/1) [11]		- High-income country [11]	- Social organization [20]	- Direct award for PPP contracts [26]
Netherlands	- High digitalized e-government (EGDI: 0.95/1) [11]		- High-income country [11]		
Germany	- High digitalized e-government (EGDI: 0.94/1) [11] - Lack of acceptance by citizens [18]	- Financial autonomy [23]	- Low revenue sources [27] - High-income country [11]		
Sweden	- Digitalized e-government (EGDI: 0.93/1) [11] - Advanced digital systems [17]	- Highest autonomy [22,18,23]	- High-income country [11]	- Small municipalities & inter-municipal-service [22,23]	
Spain	- Digitalized e-government (EGDI: 0.92/1) [11]		- High-income country [11]		
Czech Rep.	- Below average digitalization in EU [28]	- Medium local autonomy [24] - Decentralized administration [28]		- Municipal associations [24]	

Cooperation with private-sector partners (v) is also increasing across Europe in areas such as digital infrastructure, energy, waste management, and social services, even though the scope and form of PPP vary considerably across countries [26]. For instance, Nordic municipalities predominantly collaborate with private providers in innovation-oriented digital projects rather than in basic service provision [26].

4.2 International Comparison

At the global scale, differences in digital maturity are even more pronounced than in Europe due to wide disparities in economic development and institutional readiness, as shown in the concept matrix in Table 3 and in an extended version in Table 5 in the appendix. Leading digital governments are Singapore, South Korea, and Australia. Emerging economies, including India, demonstrate strong innovation potential but heterogeneous implementation across regions and local governments, reflecting disparities in administrative capacity [11]. Central governance structures (ii) can accelerate central digital e-government systems, when combined with strong administrative capability, as observed in Kazakhstan and Russia [20,29].

It should be noted that digitalization development is not as advanced as in countries with a higher degree of local autonomy. Decentralization does not necessarily mean that the development of central government platforms is less advanced. The United Kingdom demonstrates that technologically advanced central access platforms can coexist effectively with strong local autonomy [30,31]. Consistent with the European comparison, countries with a greater local autonomy tend to exhibit higher overall levels of digitalization.

Table 3 Comparison of International Countries

Country	Degree of digitalization (i)	Governance structure (ii)	Financial resources (iii)	Municipal structure (iv)	Private Co-operation (v)
Singapore	- High digitalized e-government (EGDI: 0.97/1) [11] - AI implementation [38]		- High-income country [11]		
South Korea	- High digitalized e-government (EGDI: 0.97/1) [11] - Policies & transparency for digitalization & AI implementation [4]		- High-income country [11]		
Australia	- High digitalized e-government (EGDI: 0.96/1) [11]		- High-income country [11]		
Russia	- Central e-government [32]	- Centralized, low local autonomy [29]	- 100% state funding [29]		- PPP for modernization & reducing costs [29]
USA	- Top ten e-government in 2022 based on EDGI Index [11]	- Local autonomy [30]			- 16% of public services are privatized [33]
India	- Growing e-government [35]	- Low local autonomy [30]		- Correlation between municipality size and service quality [39]	- PPP for e-government, infrastructure & services [34,35]
Iceland	- Established e-government - Inter-municipal data exchange [38]	- State-organized infrastructure [34]			- No PPP [36]
UK	- High digitalized e-government (EGDI: 0.96/1) [11] - Legal framework for digital government - AI, blockchain, IoT and 5G implementation [11]	- Local autonomy [30]	- High transparency - Low local costs [25]	- Citizen participation [25]	- PPP to maintain infrastructure & services [36]
Kazakhstan	- Over 100 digital services [28]	- Centralized - Low local autonomy [20,29]	- 100% state funding [29]	- Decision-making at national level - Low citizen participation [20]	- PPP for modernization & reducing costs [29]

Financial resources (iii) further differentiate national approaches. High-income states tend to treat digital infrastructure as a strategic public asset, whereas lower-income regions often rely on fragmented rollouts and short-term or project-based funding [11,20,31]. Evidence from Russia and Kazakhstan suggests that reliance on purely state-funded SGI is associated with lower levels of digitalization [18,29,32].

Internationally, private cooperation (v) in SGI is more widespread than in Europe. The United States outsources 16% of SGI [33]. PPP are increasingly used for the digitization and modernization of public infrastructure, as evidenced by cases in Russia, India, and the United Kingdom [29,34-36]. Although such private collaborations can accelerate innovation and scalability, transfer technical know-how, and reduce

costs, they also entail significant governance risks in contexts where oversight, accountability, and data protection frameworks are weak [29]. Technological approaches for e-government platforms vary along three dominant platform models: fully integrated one-stop systems enabling seamless access (e.g., in the UK), shared entry points linking to separate administrative systems (e.g., in Singapore), and delegated transaction models that externalize services but may compromise service quality [37]. Recently, countries such as Singapore, South Korea, and Russia have begun incorporating AI to automate user interactions and forecasting, signaling the transition toward "E-Government 3.0" [32,38].

4.3 Discussion

The analyses in Sections 4.1 and 4.2, supported by the comparative matrices illustrated in Table 2 and Table 3, provide the empirical foundation for the following patterns. By systematically comparing countries across the degree of digitalization (i), governance structures (ii), financial resources (iii), municipal structures (iv), and private cooperation (v), relationships between institutional configurations and digital outcomes become visible. Abstracting from country-specific contexts, these relationships are transferred into four patterns that explain why certain digital government and SGI initiatives succeed while others remain constrained.

The connections between (i) and (ii) reveal a pattern that underscores the *relevance of coordinated governance in driving digital progress*. High-performing countries achieve a balance between national steering and local autonomy, enabling interoperability, standardization, and innovation concurrently. This is evident in cases such as Sweden, where a high degree of municipal autonomy, combined with inter-municipal service arrangements that facilitate effective horizontal coordination among municipalities, has resulted in a relatively advanced level of digitalization [11,22,23]. The critical determinant of success is coordination capacity rather than the degree of centralization. In contrast, fragmented governance structures lacking clear alignment impede scalability and impose significant administrative burdens, thereby constraining modernization efforts.

The second pattern becomes clear when comparing (i) and (iii), emphasizing the *role of sustainable funding* and *structural capacity*. Long-term investment, local fiscal autonomy, and continuous skill development constitute necessary preconditions for effective digital transformation, as evidenced by the digitalized SGI in Denmark, where financial independence of municipalities is secured through local tax revenues [11,20]. Structural disadvantages, including fragmented financial resources, lead to territorial inequalities in the provision of smart government services, limiting both their accessibility and efficacy.

The third pattern, derived from the comparison of (i) and (v), highlights the *potential of PPP to accelerate innovation under conditions of robust regulation*. When risk allocation, transparency, and accountability are well managed, PPP can deliver technological expertise and speed, thereby enhancing the efficiency and quality of service delivery. This dynamic is illustrated by the United Kingdom, where PPP have been employed to modernize public infrastructure, including digital infrastructure [36]. Conversely, insufficient regulatory oversight may cause dependency or increase the vulnerability of public services, undermining long-term sustainability.

The fourth pattern emerges from an internal analysis of (i) and focuses on *platform-based technology models and their impact on scalability and user trust*. Unified digital platforms, providing single entry points to multiple services, enhance visibility, accessibility, and operational efficiency. For example, the United Kingdom demonstrates this through its central e-government platform, which features standardized processes and a high level of information integration, while Denmark provides comprehensive digital portals that delivers services for citizens [11,37]. Nevertheless, their effectiveness depends on alignment with governance structures and regulatory maturity; misaligned platform design can compromise the anticipated benefits despite advanced technological capabilities. The transferability of digital approaches between countries is fundamentally shaped by structural compatibility. Practices that emerge from highly coordinated governance contexts are transferable to systems with comparable administrative responsibilities, funding

mechanisms, technological maturity, and local capabilities. By contrast, in countries with fragmentation, regulatory gaps, or economic differences, solutions need to be adapted to institutional conditions. Recognizing these patterns and constraints is essential for translating successful digital practices across diverse governance environments.

5. Conclusion

This paper examined the structural and institutional factors influencing the digitalization of SGI and assessed how these factors shape the transferability of digitalization approaches across countries. Addressing a key research gap, the study moved beyond single-country perspectives by systematically comparing European and international SGI and their digitalization. Guided by the research question "*Which factors influence the implementation of SGI digitalization, and how do these factors affect the transferability of different digitalization approaches from a cross-country perspective?*", the analysis applied a systematic literature review for a European and an international comparison. This methodological approach enabled the identification of recurring patterns across governance, financial, and institutional contexts, providing a basis to transfer digitalization approaches between countries.

The analysis shows that digitalization in SGI is more shaped by governance and institutional conditions than technology. European evidence shows that strong local autonomy and municipal competence are key drivers of effective SGI digitalization. At the same time, findings on global SGI confirm that local autonomy and competence alone do not determine digital performance, since institutional coordination is equally important. Sustainable funding and financial capacities are equally critical, as long-term investment and fiscal autonomy determine the quality of digital SGI. PPP can enhance efficiency and innovation if aligned with regulatory and structural frameworks, though misalignment risks dependency, accountability, and data protection issues. Platform-based digital solutions enhance scalability and user trust if aligned with governance structures and institutional capacities. Regarding transferability, the study demonstrates that digitalization approaches cannot be directly replicated across countries. Consequently, best practices require contextual adaptation rather than simple adoption. The results confirm the relevance of the analytical dimensions (ii) to (v) identified in the related background.

Overall, this paper contributes to a more nuanced understanding of SGI digitalization by identifying transferable patterns and structural constraints. The results support policymakers, technology providers, and public managers in designing resilient, context-sensitive digital transformation strategies by emphasizing the importance of governance structures, funding mechanisms, and private-sector collaboration. However, the findings of this study were subject to several limitations. The analysis primarily addressed e-government rather than SGI more broadly, with limited consideration of AI applications, as they remain underrepresented in the existing comparative literature on SGI. In addition, the availability and depth of comparable literature vary across countries, which constrains the generalizability of cross-country insights. The findings reflect a current excerpt of state-of-the-art literature addressing cross-country comparisons of SGI obtained from a systematic literature review. Future research could complement these findings by using alternative methodologies, such as case studies or expert interviews, to capture practice-oriented dynamics and to deepen the empirical basis for evaluating the transferability of digitalization approaches.

Acknowledgements

This research is co-funded by the European Union and the Ministry for Culture and Science of the State of North Rhine-Westphalia (MKW) within the “Innovation Competition NEXT.IN.NRW” (funding number: EFRE-20800508) and managed by the Projektträger Jülich (PTJ). The authors are responsible for the content of this publication.

Appendix

Table 4 Extended Comparison of European Countries

Country	Degree of digitalization (i)	Governance structure (ii)	Financial resources (iii)	Municipal structure (iv)	Private Co-operation (v)
Denmark	- Highest digitalized e-government (EGDI: 0.99/1) [11] - Policies and transparency mechanism for digitalization [11] - Comprehensive digital portals for citizens, businesses, and health services through platforms with cross-sector collaboration [11] - Sustainability in public procurement and explores AI, robotics, and 5G infrastructure to enhance public services [4]	- Decentralized - Local autonomy [20]	- Financial independence of municipalities through tax revenues [20] - High-income country [11] - Local spending on social protection, education, and public safety [25]	- Citizen-oriented urban and infrastructure policy, sustainable and social design [20]	- Public service obligations are commissioned by the state or local authorities [26]
Estonia	- High digitalized e-government (EGDI: 0.97/1) [11] - Policies and transparency mechanism for digitalization [11] - Robust infrastructure and forward-thinking initiatives [11] - Emphasizes the importance of proactive services, ensuring digital accessibility and inclusion across all demographic groups by 2030 [11] - Comprehensive digital identity system, enabling seamless online access to public services [11] - Online services, AI implementation [4]	- High degree of local autonomy [20] - Centralization of fire and rescue services [22]	- High-income country [11]	- Mandatory inclusion of citizens in decision-making [20]	- Partnerships provided by cities, restructuring to public tenders, - Public service obligation (PSO) contracts are awarded directly, with only the state being responsible for awarding services [26]
Finland	- High digitalized e-government (EGDI: 0.96/1) [11]		- High-income country [11]	- Focus on social organization of local communities [20]	- Public service obligation (PSO) contracts are awarded directly, with only the state being responsible for awarding services [26]
Netherlands	- High digitalized e-government (EGDI: 0.95/1) [11]		- High-income country [11]		
Germany	- High digitalized e-government (EGDI: 0.94/1) [11] - Lack of acceptance by citizens [18]	- High financial autonomy, local government systems for social welfare provision [23]	- Low municipal revenue sources (taxation on real estate and commercial goods)[27] - High-income country [11]		
Sweden	- Digitalized e-government (EGDI: 0.93/1) [11] - Advanced digital systems [18] - Multilingual e-government services [18]	- Highest fiscal autonomy, strong local government systems for social welfare provision [22,18,23]	- High-income country [11] - Local government expenditure accounts for 24% of GDP [21], primarily in the areas of health, social services, and education [25] -high local expenditure on health (7% of GDP), social protection (6.7% of GDP) and	- Small number of large municipalities with high population [22] - 290 municipalities that provide municipal or inter-municipal services; independent provision of water and energy supplies [23]	

			education (5.2% of GDP) [25]		
Spain	- Digitalized e-government (EGDI: 0.92/1) [11]		- High-income country [11]		
Czech Rep.	- Below average digitalization in EU [28]	- Medium local autonomy (Local autonomy index 64.4/100) [24] - Decentralized administration [28]		- Associations in voluntary municipal associations and action groups [24]	

Table 5 Extended Comparison of International Countries

Country	Degree of digitalization (i)	Governance structure (ii)	Financial resources (iii)	Municipal structure (iv)	Private Co-operation (v)
Singapore	- High digitalized e-government (EGDI: 0.97/1) [11] - AI implementation as a central driver [38]		- High-income country [11]		
South Korea	- Second-highest digitalized e-government (EGDI: 0.97/1) [11] - Policies & transparency for digitalization & AI implementation [4]		- High-income country [11]		
Australia	- High digitalized e-government (EGDI: 0.96/1) [11]		- High-income country [11]		
Russia	- Central e-government [32]	- Centralized [29] - Low local autonomy with little decision-making power at the local level [29]	- 100% state funding [29]		- Cooperation with PPP to modernize public infrastructure, utilize private expertise, and relieve pressure on state budgets in the areas of transport, energy and water supply, waste management, and social service [29]
USA	- Top ten e-government in 2022 based on EDGI Index, but not listed in the top performing countries in 2024 [11]	- Local autonomy [30]			- 16% of public services are privatized [33] - Health Care fully privatized [31]
India	- Growing e-government [36]	- Low local autonomy [30] - Water supply managed by local authorities [39]		- Correlation between municipality size and service quality [39] - Slums have limited access to basic services, and rural areas have poorer provision [39]	- Development and provision of e-government services through PPP [35] - For public infrastructure and services in the areas of transport, energy and water supply, education and health, 1,540 projects [34]
Iceland	- Established e-government - Inter-municipal data exchange between administrative areas [38]	- State organized infrastructure, e.g., energy supply [36]			- No cooperation with private sector actors [36] - Publicly controlled structures in energy supply [36]
UK	- Third highest digitalized e-government (EGDI: 0.96/1) [11] - Local implementation strategies for the digital provision of services [40]	- High degree of local autonomy [30] - Municipal decision-making on care services [31]	- High transparency [25] - Low local costs on health, high local public costs	- Citizen participation [25]	- Cooperation with PPP to maintain public infrastructure and services in the

	- Central platform for e-government, standardized, and a high level of information integration [37] - Legal framework for digital government - AI, blockchain, IoT, and 5G implementation [11]		on education [25] - Financial restrictions on care services [30] - Mandatory citizen participation in budgeting [20] - Flexible municipal administrative structures [20]		areas of transport, energy and water supply, education and healthcare; 1570 projects [34] - Market-oriented energy supply by private service providers [36]
Kazakhstan	- Over 100 digital services [28]	- Centralized - Low financial and administrative local autonomy [20,29]	- 100% state funding [29] - Low influence of local tax revenues [20]	- Decision-making at the national level - Low citizen participation [20] - Decision-making at the national level [20]	- Cooperation with PPP to modernize public infrastructure, utilities private expertise and relieve state budgets in the areas of transport, energy and water supply, waste management, and social facilities [29]

References

[1] European Economic and Social Committee, 2012. Services of general interest. https://www.eesc.europa.eu/en/policies/policy-areas/services-general-interest. Accessed 14 December 2025.

[2] Karamchand, P., 2021. Examining the Challenges and Opportunities of Managing Public Services in a Complex and Interconnected World. Journal of Public Representative and Society Provision (JPRSP) 1 (2), pp. 51–56.

[3] Peng, Y., Ahmad, S.F., Irshad, M., Al-Razgan, M., Ali, Y.A., Awwad, E.M., 2023. Impact of Digitalization on Process Optimization and Decision-Making towards Sustainability: The Moderating Role of Environmental Regulation. Sustainability 15 (20), p. 15156.

[4] OECD, 2025. Government at a Glance 2025. OECD Publishing, Paris.

[5] Eichholz, L., 2025. Municipal AI integration: a structured approach. Frontiers of Urban and Rural Planning (Front. Urban Rural Plan.) 3 (1).

[6] Kaelan Boyd, 2024. Local Government Practitioners Weigh in on AI. https://icma.org/articles/article/local-government-practitioners-weigh-ai. Accessed 14 December 2025.

[7] Bloomberg Philanthropics, 2023, October 18. State of Cities: Generative AI in Local Government. Bloomberg Philantropics.

[8] Große, N., Hefft, D., Wolny, M., Eichholz, J., 2025. Development Of A Taxonomy On Services Of General Interest, in: D. Herberger & M. Hübner (Ed.), Proceedings of the 7th Conference on Production Systems and Logistics, Hannover: Institutionelles Repositorium der Leibniz Universität Hannover, pp. 172-182.

[9] Wollmann, H., Koprić, I., Marcou, G., 2016. Public and social services in Europe: From public and municipal to private sector provision. Palgrave Macmillan, Basingstoke, 342 pp.

[10] Mergel, I., Edelmann, N., Haug, N., 2019. Defining digital transformation: Results from expert interviews. Government Information Quarterly 36 (4), p. 101385.

[11] United Nations Department of Economic and Social Affairs, 2024. UN E-Government Survey 2024. United Nations.

[12] Almquist, R., Grossi, G., van Helden, G.J., Reichard, C., 2013. Public sector governance and accountability. Critical Perspectives on Accounting 24 (7-8), pp. 479–487.

[13] Hodge, G.A., Greve, C., 2019. The logic of public-private partnerships: The enduring interdependency of politics and markets. Edward Elgar Publishing, Cheltenham, UK, Northampton, MA, 293 pp.

[14] vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. http://www.alexandria.unisg.ch/Publikationen/67910.

[15] Cooper, H.M., 1988. Organizing Knowledge Syntheses: A Taxonomy of Literature Reviews. Knowledge in Society 1 (1),, pp. 104–126.

[16] Webster, J., Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly, Jun., 2002, Vol. 26, No. 2 (Jun., 2002),

[17] Wauters, P., 2006. Benchmarking e-government policy within the e-Europe programme. Aslib Proceedings 58 (5), pp. 389–403.

[18] Akkaya, C., Wolf, P., Krcmar, H., 2012. The Surprisingly Low Effect of National Culture on E-Government Adoption: A Cross-Cultural Comparison. Americas Conference on Information Systems (AMCIS) 2012 Proceedings. 1.

[19] Pisar, P., PRIŠČÁKOVÁ, S., ŠPAČEK, D., NEMEC, J., 2022. Digitization as a tool of E-Goverment in selected public services: International comparison of Slovakia and the Chech Republik. ADMINISTRATIE SI MANAGEMENT PUBLIC (AMP) 1 (39), pp. 111–132.

[20] Abdullayev, A., Kadyrova, M., Ydyrys, S., 2025. Modernization of the Local Government System: Development Prospects and Implementation Mechanisms. Emerg Sci J 9 (3), pp. 1434–1476.

[21] Puolokainen, T., 2017. Reforming fire and rescue services: a comparative study of Estonia and Georgia. International Journal of Public Sector Management (IJPSM) 30 (3), pp. 227–240.

[22] BAUBY, P., 2012. Local services of general economic interest in Europe. Water sevice: What are the challenges? Ann Public Coop Econ 83 (4), pp. 561–583.

[23] Kuhlmann, S., Dumas, B.P., Heuberger, M., 2022. Introduction, in: Kuhlmann, S., Dumas, B.P., Heuberger, M. (Eds.), The Capacity of Local Governments in Europe. Springer International Publishing, Cham, pp. 1–6.

[24] Bel, G., Bischoff, I., Blåka, S., Casula, M., Lysek, J., Swianiewicz, P., Tavares, A.F., Voorn, B., 2023. Styles of inter-municipal cooperation and the multiple principal problem: a comparative analysis of European Economic Area countries. Local Government Studies 49 (2), pp. 422–445.

[25] Halásková, M., Halásková, R., Bednář, P., 2020. Public Services and their Financial Allocation in the European Context. Scientific Papers of the University of Pardubice, Series D: Faculty of Economics and Administration (SciPap) 28 (2).

[26] Dolinayova, A., Cerna, L., Domeny, I., 2022. Competition on the Domestic Rail Passenger Transport Market Under Public Service Obligation in Some Selected European Countries and Slovak Republic, in: Marinov, M., Piip, J. (Eds.), Sustainable Rail Transport 4. Springer International Publishing, Cham, pp. 269–323.

[27] Leisink, P., Bach, S., 2014. Economic crisis and municipal public service employment: comparing developments in seven EU Member States. Transfer: European Review of Labour and Research 20 (3), pp. 327–342.

[28] Ulman, M., Ualiyev, N.S., Toregozhina, M.B., 2016. Do Digital Public Services Matter? A Comparative Study of the Czech Republic and the Republic of Kazakhstan. AOL 08 (02), pp. 121–133.

[29] Mouraviev, N., Kakabadse, N., Robinson, I., 2012. Concessionary Nature of Public-Private Partnerships in Russia and Kazakhstan: A Critical Review. International Journal of Public Administration 35 (6), pp. 410–420.

[30] Önder, M., Nyadera, I.N., Islam, M.N., 2022. Cross-Natıonal Comparıson of Publıc Admınıstratıon Systems: Selected Cases Across the Globe, in: Önder, M., Nyadera, I.N., Islam, M.N. (Eds.), The Palgrave Handbook of Comparative Public Administration. Springer Nature Singapore, Singapore, pp. 743–761.

[31] Mos, P., Reckers-Droog, V., 2024. Examining the underpinnings of decisions to allocate public resources to social care: A systematic review. Social science & medicine (1982) 355, p. 117093.

[32] Demidova, S., Kulachinskaya, A., Razletovskaia, V., Svetlichnyy, S., 2024. Development Trends of E-government and Public Finance Management in the Context of Digitalisation, in: Sari, M., Kulachinskaya, A. (Eds.), Digital Transformation: What are the Smart Cities Today?, vol. 846. Springer Nature Switzerland, Cham, pp. 161–172.

[33] Hebdon, R., Jalette, P., 2008. The Restructuring of Municipal Services: A Canada—United States Comparison. Environment and Planning C: Government and Policy (Environ Plann C Gov Policy) 26 (1), pp. 144–158.

[34] Cheng, Z., Ke, Y., Yang, Z., Cai, J., Wang, H., 2020. Diversification or convergence. ECAM 27 (6), pp. 1315–1335.

[35] Pardhasaradhi, Y., Ahmed, S., 2007. Efficiency of electronic public service delivery in India, in: Proceedings of the 1st international conference on Theory and practice of electronic governance. ICEGOV '07: International conference on theory and practice in electronic governance, Macao China. 10 12 2007 13 12 2007. ACM, New York, NY, USA, pp. 357–365.

[36] Allemand, R., Dreyfus, M., Magnússon, M.Á.S., McEldowney, J. Local Government and the Energy Sector: A Comparison of France, Iceland and the United Kingdom, in: , pp. 233–247.

[37] T., Kohlborn, Thomas, S., Weiss, Simon, J., Poeppelbuss, Jens, A., Korthaus, Axel, E., Fielt, Erwin, 2010. Online service delivery models - An international comparison in the public sector. ACIS 2010 Proceedings - 21st Australasian Conference on Information Systems.

[38] Hadi, H., Elhassani, I., Sekkat, S., 2021. Electronic Public Services in the AI Era, in: Ben Ahmed, M., Rakıp Karaş, İ., Santos, D., Sergeyeva, O., Boudhir, A.A. (Eds.), Innovations in Smart Cities Applications Volume 4, vol. 183. Springer International Publishing, Cham, pp. 70–82.

[39] Sahasranaman, A., Bettencourt, L.M., 2021. Life between the city and the village: Scaling analysis of service access in Indian urban slums. World Development 142, p. 105435.

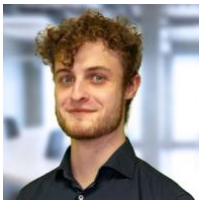
[40] Weerakkody, V., El-Haddadeh, R., Sabol, T., Ghoneim, A., Dzupka, P., 2012. E-government implementation strategies in developed and transition economies: A comparative study. International Journal of Information Management 32 (1), pp. 66–74.

Biography

Eva Paulina Ziehm (*1999) is a research assistant at the Chair of Enterprise Logistics, TU Dortmund University since 2025 and works in the department of factory planning, maintenance planning, and industrial services. Her research focuses data management with in industrial services.

Dr. Nick Große (*1992) has been a postdoctoral researcher at TU Dortmund University since 2022. As a research coordinator at the Chair of Enterprise Logistics, he advises doctoral researchers in the fields of factory planning, maintenance planning, and industrial services. His research focuses on mechanisms that establish and maintain trust in industrial ecosystems, applying innovative key technologies.

Jonas Eichholz (*1996) is a research assistant at the Chair of Enterprise Logistics, TU Dortmund University since 2020 and works in the department factory planning, maintenance planning, and industrial services. His research focuses on co-creation, services and data management with a focus on industrial services.



Leon Jungh (*1999) has been working as a student assistant at the Chair of Enterprise Logistics, TU Dortmund University since 2024. He is a master's student in Industrial Engineering at TU Dortmund.

Marius Klinsmer Tchapga Leumen (*1997) has been working as a student research assistant at the Chair of Enterprise Logistics, TU Dortmund University since 2024. He is a master's student in Mechanical Engineering at TU Dortmund University, specializing in production engineering.

8th Conference on Production Systems and Logistics

Implicit Flexibility in Agent Behavior within Cloud Manufacturing

Robin Huwa, Marcus Albrecht, Andreas Leinenbach, Johannes Schilp
University of Augsburg, Faculty of Applied Computer Science, Chair of Digital Manufacturing, Augsburg, Germany

Abstract

The concept of cloud manufacturing leads to a significant paradigm change within the industrial context: Many technical hurdles are being mastered, mostly based on digitalization and networking of production sites. Furthermore, production planning and control must evolve in the management dimension in order to address new kinds of concurrent yet independent production processes. However, centralized or hierarchical production planning is inadequate for the different participating entities with individual, diverging goals that interact within and across production sites.

In this work, we propose that agent-based modeling more accurately reflects entity behavior intricacies and interrelated interactions within a dynamic yet structured environment. This environment comprises individual production sites with a range of capabilities based on their internal structure and transportation logistics between those sites. Production agents utilize these networked sites for their own production while protecting their intellectual property from other agents by limiting information exchange with other agents, thereby forming a multi-agent system. Hence, the agent-based modeling technique can facilitate modeling and predicting agent performance. One major aspect of this modeling technique in the context of cloud manufacturing is the implicitly defined (in-)flexibility of an agent, which greatly influences its performance within the simulated dynamic environment. This implicit flexibility within agent behavior, as well as its boundaries, are analyzed and quantified within this work.

Keywords

Agent-Based Modeling; Industry 4.0; Flexibility; Decentralized Resource Scheduling; Multi-Agent System; Production System

1. Introduction

Cloud Manufacturing (CMfg) has been a dedicated research field since the advent of the fourth Industrial Revolution, and continues to be in the fifth Industrial Revolution. The main research and development impulse comes from the number of benefits being proposed as the result of this new approach for the interconnected organization of production systems within value networks. These benefits include a greater degree of overall equipment efficiency, efficient sharing of resources, more flexible production flow, production optimization across distributed locations, and many more [1–4].

However, due to CMfg being such a radical paradigm change, different challenges have to be met. These include various legal, managerial and organizational challenges, as well as many technical hurdles that remain to be mastered. The latter contain for example the differing levels of digitalization and networking between production systems. Furthermore, from the perspective of production system management and planning, CMfg leads to less centralized, unified control over actively scheduled production tasks, because of the decoupling of product owner and production resource owner. This is due to the additional external,

DOI: https://doi.org/10.15488/20993
ISSN: 2701-6277

potentially weakly predictable production tasks which are introduced into the local, known production mix. Hence, the question arises: How to predict the behavior of external participants within local production system from the global CMfg value network point of view? We propose the utilization of Agent-Based Modeling (ABM) to model product owners as individual agents which plan and execute their production processes within an the value network environment comprising interconnected production systems. Our approach enhances intellectual property protection between these decentralized agents by minimizing necessary scheduling consensus interactions, as well as allowing for extensive individuality and flexibility for the implementation of the behavior of these agents.

2. Background

While different centralized approaches to model CMfg are being researched and evaluated in practice, we propose that decentralized methodologies such as agent-based modeling more appropriately reflect the intricacies of CMfg. Additionally, the concept of quantitative flexibility is enabled by such methodologies, all of which we will explore further.

2.1 Cloud Manufacturing Approaches

CMfg as a concept comprises a number of aspects dealing with the optimization of production across the whole value network. At its core, its goal is to loosen the traditionally rigid and hierarchical organization of production processes: Instead of fixed production flow between entities in a value network, these decouple their manufacturing capabilities and offer them individually as services, leading to Equipment-as-a-Service [5] and Manufacturing-as-a-Service [6]. While in practice this concept is yet to be properly implemented, in literature various definitions as well as working models can be found: "ubiquitous, hassle-free, and on-demand network access to a shared pool of manufacturing resources" [1], "[CMfg] enables companies to instantly access and share resources from anywhere worldwide, including data, hardware, and software" [2], "[CMfg] represents an emerging paradigm where distributed resources are encapsulated into cloud services and centrally managed. This network of shared resources enables customers to access on-demand services supporting the entire product lifecycle." [7]

Overall, CMfg is treated in literature as a scheduling problem which is more complex due to the internal heterogeneity of individual entities as well as the larger scale beyond single production processes but across value networks. The approaches in literature to solve this problem yet remain, as with previous solutions beforehand, centralized simulations and models. We propose that decentralized modeling such as ABM is more appropriate for such a heterogeneous, decentralized planning and scheduling problem.

2.2 Agent-Based Modeling

ABM is a modeling technique to manage heterogeneous entities, called agents, which interact within an environment to reach their goals [8]. The aspect of heterogeneity is herein central: To fulfill their individual goals, agents follow individual strategies and take independent actions in the environment. Generic examples in the literature include: The behavior of commuters in urban transportation [9, 10], The socio-economical decision making effects of buyers in markets [11], or the complex interactions of microbial communities [12]. Hence, similar to population-based meta heuristics, ABM is appropriate for many differing research fields, from physics-based interactions to social networks.

The main types of settings which ABM is useful for revolve around two main modeling characteristics: Heterogeneity and decentralization [8]. This can also be seen in the aforementioned examples. As heterogeneous entities in ABM take individual actions according to some sort of internal strategy, which by design are not (necessarily) coordinated with any other agents, they exhibit these two characteristics.

While traditionally in the industrial context most analyses and modeling techniques focus on centralized simulation and optimization with one central, core algorithm for goal fulfillment, we propose that for the paradigm change induced by CMfg a decentralized simulation model such as ABM is more appropriate.

2.3 Flexibility in the Industrial Context

In the industrial context, the term flexibility is used often, especially in recent years, and is attributed to highly diverse aspects within that context. Prominent examples include additive manufacturing techniques [13] or flexible manufacturing systems [14]. Our previous research has shown that the term flexibility is applied to many different aspects of manufacturing and many methodological approaches aim to increase flexibility in production, for products, processes and resources, yet do not define the improvement numerically.

For the quantification of flexibility, we have previously focused on individual dimensions of flexibility. To that end, we categorized flexibility into three main groups following the Process-Product-Resource Model [15]. As examples of process flexibility, we have defined and evaluated Temporal Flexibility [16] and are currently researching Operations Order Flexibility, both of which compute the feasible solution space for the deviation of active production processes without impacting other processes. In this work, we pose that similar implicit flexibilities are found in agent behavior within ABM, which deserves further quantitative research.

3. Decentralized Agents for Cloud Manufacturing

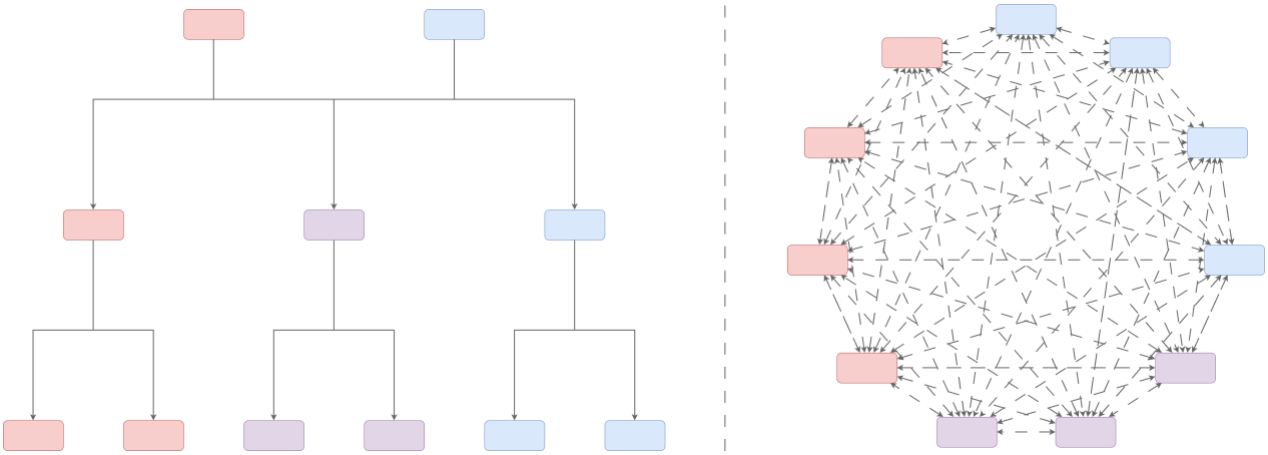


Figure 1: Comparison of traditionally hierarchical organization of production systems within a value network (left) to the less rigidly structured, open organization in CMfg (right).

The paradigm change towards CMfg induces a restructuring and reorganization of production processes across production systems within value networks. Traditionally, production systems within a value network are structured hierarchically, as shown in Figure 1: Production flows through multiple supplying production systems and generally flows together the higher in the hierarchy it is, until a completed product is the result of the final production system. Production systems may be part of multiple production hierarchies, leading to parallel structures. These structures so far are optimized for mass production with a fixed product mix, hence they traditionally are rigid. Of note is the fact that the producing entity on top of the production hierarchy is also the product owner, leading to a tight coupling between product goals and production process goals.

While CMfg does not change the fact that production flow across the value network necessarily converges towards a single production system on top of the hierarchy, it aims to enable more flexible production flow by reducing the rigidity of production flow. For example, in CMfg the production system is decoupled

from the product: A product owner defines the parameters for production of its product, by design fairly independent from the existing production systems. This production demand is then offered to the value network, wherein the production process from start to finished product has to be decided by the product owner. The owner is hereby taking multiple optimization metrics into account, mainly time and cost. However, this owner is decoupled from any individual production system, hence has less limitations on and inhibitions for product design.

Especially with related platform-based concepts such as Equipment-as-a-Service [5] and Manufacturing-as-a-Service [6], overall production process creation and definition tends to become less rigid, meaning more flexible but also more chaotic. As product owners access and plan utilizing independent production systems to receive the most optimal production flow, from the perspective of the production systems owners across the value network many external, less predictable production demands are put onto the production system. This unpredictability of external, decentralized influences on local production puts a strain on the plannability and optimality of tasks.

Approaches for scheduling traditional production processes in value networks are in most cases based on centralized control of all production assets in question at the head of the production hierarchy. Examples include centralized AI-driven methodologies [17–19], but also classical evolutionary algorithms [2, 4, 20, 21] and traditional rules-based approaches [1, 22]. However, with the previously described decentralization in CMfg these centralized approaches become less applicable. Instead of centralized task scheduling, individual entities within the value network access distributed resources for their independent production tasks.

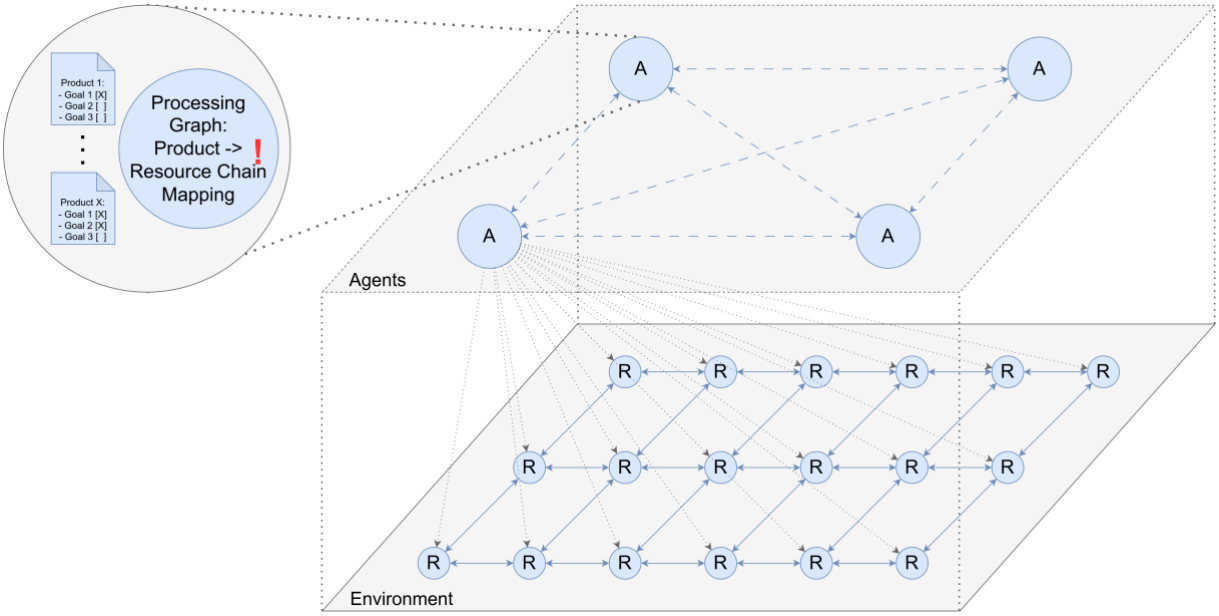


Figure 2: Depiction of the agent - environment split in ABM for CMfg, whereas agents (A) manage products and their production processes, and interact with an environment consisting of production resources (R) of distributed, interconnected production systems within a value network.

We propose that decentralized modeling is more appropriate for this distributed production planning & scheduling problem [7]. In the context of ABM, product owners of CMfg become product agents which manage the production process and other details concerning the product. Opposite of that, the production systems within the value network are modeled as the environment of the aforementioned agents, so that agents interact with the environment to fulfill their production goals. In that way, the interrelations between agents and the environment can be predicted.

The environment, as shown in Figure 2, consists of interconnected production resources. These model the production capabilities of production systems within a value network, which are connected via logistics

capabilities. These capabilities vary greatly between value systems and industrial fields in their details, inputs and outputs as well as associated costs and other metadata. These capabilities are offered up to the product agents to utilize for their production targets.

In general, agents act freely within the modeled environment and are logically decoupled from it, as shown in Figure 2. They can choose from the available production and logistics resources of the individual production systems within the value network to create their products. For that, the agents base their actions on a detailed description of the product, containing the necessary production steps as well as the desired final state of it.

The main, active task managed by the product agent is called the processing graph, which is highlighted in Figure 2: Firstly, in the planning stage, a feasible processing graph which covers the product goals is determined based on the available resources within the environment. This focuses on the proper order and chaining of individual resources, which becomes especially complex when resources between individual resources are involved due to added logistics. Secondly, in the production stage, the agent ensures the follow-through of the processing graph and reacts to externally induced changes to the environments, such as added or removed resources, by adapting the processing graph according to its goals. What these goals are, and what strategies are employed to achieve them, is to be defined by each agent individually for both planning and production phases.

With this modeling approach, following beneficial effects are induced:

- Agents can employ greatly different strategies during the planning and production phases to manage the processing graph, which more accurately reflects the expectable heterogeneity of product agents and product owners.
- As agents interact only with the environment and not necessarily with each other, knowledge exchange or information loss between agents is minimal. This ensures the protection of intellectual property about the product and the processing graph of each agent.
- The environment can evolve dynamically during the simulation of production due to external changes, which reflects conditions of real production more closely. Agents have to be able to react to these changes.

Based on these findings, we conclude that ABM is a good match for modeling the decentralized intricacies of CMfg, hence can be employed to predict entity behavior within this paradigm.

4. Exposing Implicit Flexibility Encoded in Agent Behavior

As flexibility has become an overarching attribute of many of the approaches developed within the contexts of both the Industry 4.0 and 5.0, it is worth evaluating the flexibility of ABM for CMfg. Our previous research has shown a pervasive tendency in literature that the terminology surrounding flexibility remains not strictly defined, hence is based on the implicit understanding [16]. Similarly, ABM for CMfg may be described as more flexible in comparison to traditionally centralized approaches: The restraint for homogeneity of participating entities is removed, product and resource owners are decoupled, production flow hierarchies are relaxed. Furthermore, the individual strategies of agents may exhibit more or less flexibility depending on their design. We propose that kind of flexibility dimensions of agents are quantifiable, hence are algorithmically usable.

As the main focus of product agents are the product and process categories of the Product-Process-Resources Model [23], flexibility sub-dimensions of these two categories are applicable as based on [15]. This leads to concepts such as Temporal Flexibility [16] being applicable. In this work, we will present one flexibility dimension per category, as well as their effects in the planning and production phase of the modeled agents: Capability Matching Flexibility and Planning Horizon Flexibility.

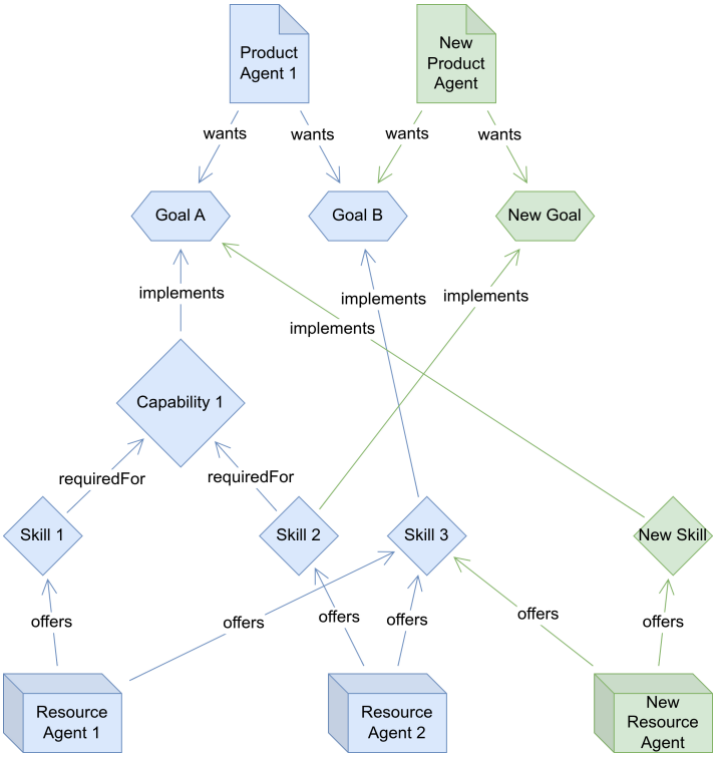


Figure 3: The depiction of the necessary mapping between resource capability offers and product goal demands to enable production, from our previous work [24].

Capability Matching Flexibility describes the ability of the production process to utilize alternative resources to reach the same product goal. While the processing graph of an agent comprises an ordered set of production steps, which utilize known resources to reach the final state of the product, this flexibility allows for alternative steps to be used: A prominent example is the usage of additive manufacturing instead of the CNC-machining or molding of materials [13]. This usually leads to a subset of X original steps of the processing graph to be replaced by another count of Y new ones. This is depicted in Figure 3. This kind of flexibility is useful in both the planning phase and the production phase, allowing for both static optimization before production even begins but also quick adaptation of the active processing graph during production in case of the addition or removal of resources. Hence, new and better resources may be employed when made available, or greater delays in production may be avoided as planned resources are unavailable due to any reason.

Alternative capabilities by default not only differ in their execution during production, but also in their results: Quality, cost, time or energy consumption are only some of the possible differing metadata dimensions where the alternative processing steps Y can differ from the original ones X. To quantify the maximum allowable difference between the original steps and any alternative, a capability difference function $f_{capdiff}(X,Y)$ which multi-dimensionally computes the differences between X and Y may be defined. And each agent determines its maximum allowable capability difference vector $l_{maxcapdiff}$ so that for any alternative processing graph steps Y compared to X along the capability difference function $f_{capdiff}(X,Y)$ remains less than $l_{maxcapdiff}$. This is one quick example how capability differences may be quantified, however, the determination of $l_{maxcapdiff}$ for each agent is highly individual and depends on the definition of the desired final product state.

Planning Horizon Flexibility describes the ability of the production process to quickly adapt its processing graph to dynamic events. Hence, this kind of flexibility is only applicable during the production phase, when the processing graph is being actively followed: If any external event induces a change in the

topology of the environment, this kind of flexibility quantifies how much time is necessary to adapt the planned processing graph to these new environmental circumstances and to deviate from the original processing graph to the newly adjusted one. External events may be both positive and negative to the production goals of the agent; The addition, removal or adaptation of resources in the environment may lead to either positive or negative results for the agent. In both cases it is important for the agent to react to these changes, hence a high amount of planning horizon flexibility leads to a quick adaptation to the new circumstances.

How quickly an agent has to react depends on the severity of the impact of the change of the environment. Technically, for positive changes, no direct reaction by the agent is necessary: Addition or improvement of resources leads to no issues with following the planned processing graph. However, to reach optimality in production, re-planning and -scheduling may become necessary. For negative changes, such as the removal of resources which where included in the processing graph, re-planning becomes highly important. Two dimensions describe the need for high planning horizon flexibility: impact severity and temporal urgency.

To quantify this dimension of agent flexibility, firstly the default planned state of the processing graph has to be defined: The next N steps of the processing graph are scheduled on various resources in the environment. The flexibility hereby can be defined as follows: $N_{flex} \leq N$, so that only the next $N - N_{flex} = N_{necessary}$ steps have to be finished definitely before proper re-planning can occur. With any $N_{necessary} > 0$, the issue remains with maximally severe or time-sensitive impacts on the processing graph: These cannot be managed without interruptions to production. Yet in any other situation, this flexibility might minimize any negative effect.

All in all, these flexibility dimensions show that agents in ABM can include implicit or explicit flexibility in their internal production strategies and behavior in the context of CMfg. Furthermore, agent flexibility can be defined more clearly and quantified accurately, as shown with the previous two examples. This enables the usage and evaluation of flexibility metrics in simulations, so that the effects of available flexibility can be compared against its drawbacks in a quantitative manner.

5. Conclusion

In this work, we discussed the concept of CMfg from an architectural perspective and what modeling approaches are currently being researched. Our main contribution in this discussion is the proposition that decentralized modeling approaches, in particular ABM, are more applicable due the matching of organizational structures between ABM and CMfg. Furthermore, we deliberated upon the concept of flexibility in ABM for CMfg and how agent behavior can exhibit various kinds of flexibility, implicitly or explicitly. For the latter, we defined two dimensions of agent flexibility explicitly and quantifiably, which enables their usage in further evaluation of the impact of flexibility in CMfg.

Our research shows that due to the paradigm change induced by CMfg, also the analytical and simulative approaches to model future production have to adapt and evolve in their architecture: The decentralization due to CMfg necessitates a decentralized modeling of the planning and scheduling problem. Furthermore, the paradigm change induces enhancements in the flexibility of the entities in CMfg as well as for the overall system, which this work explored and which deserves further research.

As future work, the implementation task of modeling CMfg with ABM remains: This includes properly modeling a realistic environment based on existing value networks, their production systems and other entities, as well as modeling agent behavior based on real-life decision making processes. The evaluation of such a simulation is going to result in greater understanding of CMfg as well as the impacts of various changes in the intricacies in agent behavior.

Acknowledgments

This paper has been created in the context of the AI-Production-Network, which in turn has been enabled and supported by the Hitech Agenda of the Bavarian State Ministry of Education, Science and the Arts.

References

[1] Gharibvand V, Kolamroudi MK, Zeeshan Q, Çınar ZM, Sahmani S, Asmael M, Safaei B. Cloud based manufacturing: A review of recent developments in architectures, technologies, infrastructures, platforms and associated challenges. The International Journal of Advanced Manufacturing Technology. 2024;131(1):93–123. doi:10.1007/s00170-024-12989-y

[2] Burra LR, Bonam J, Sharma YK, Abirami A, Devi G, Rani U, Madavi S. Optimizing cloud manufacturing service composition using social group optimization. International Journal on Interactive Design and Manufacturing (IJIDeM). 2025 Sept 2 [accessed 2025 Nov 27]. https://doi.org/10.1007/s12008-025-02375-7. doi:10.1007/s12008-025-02375-7

[3] Kavre MS, Sunnapwar VK, Gardas BB. Analysing the role of cloud manufacturing for achieving sustainable business performance in the era of digital economy. International Journal of Organizational Analysis. 2025 Aug 12 [accessed 2025 Nov 27]. https://doi.org/10.1108/IJOA-04-2025-5424. doi:10.1108/IJOA-04-2025-5424

[4] Liu S, Deng Q, Liu X, Luo Q, Li F, Jiang C. Dual-service integrated scheduling of manufacturing and logistics for multiple tasks in cloud manufacturing. Expert Systems with Applications. 2024;235:121129. doi:10.1016/j.eswa.2023.121129

[5] Stojkovski I, Achleitner A-K, Lange T. Equipment as a Service: The Transition Towards Usage-Based Business Models. 2021 [accessed 2025 Dec 12]. https://papers.ssrn.com/abstract=3763004. doi:10.2139/ssrn.3763004

[6] Karamanli A, Xanthopoulos A, Gasteratos A, Koulouriotis D. Manufacturing-as-a-Service: A Systematic Review of the Literature. In: Kostavelis I, Folinas D, Aidonis D, Achillas C, editors. Supply Chains. Cham: Springer Nature Switzerland; 2025. p. 269–281. doi:10.1007/978-3-031-69344-1_20

[7] Renna P. A Review of Game Theory Models to Support Production Planning, Scheduling, Cloud Manufacturing and Sustainable Production Systems. Designs. 2024;8(2):26. doi:10.3390/designs8020026

[8] Collier N, Ozik J. Distributed Agent-Based Simulation with Repast4Py. In: 2022 Winter Simulation Conference (WSC). 2022. p. 192–206. https://ieeexplore.ieee.org/abstract/document/10015389. doi:10.1109/WSC57314.2022.10015389

[9] Bastarianto FF, Hancock TO, Choudhury CF, Manley E. Agent-based models in urban transportation: review, challenges, and opportunities. European Transport Research Review. 2023;15(1):19. doi:10.1186/s12544-023-00590-5

[10] Huang J, Cui Y, Zhang L, Tong W, Shi Y, Liu Z. An Overview of Agent-Based Models for Transport Simulation and Analysis. Journal of Advanced Transportation. 2022;2022(1):1252534. doi:10.1155/2022/1252534

[11] Axtell RL, Farmer JD. Agent-Based Modeling in Economics and Finance: Past, Present, and Future. Journal of Economic Literature. 2025;63(1):197–287. doi:10.1257/jel.20221319

[12] Nagarajan K, Ni C, Lu T. Agent-Based Modeling of Microbial Communities. ACS Synthetic Biology. 2022;11(11):3564–3574. doi:10.1021/acssynbio.2c00411

[13] Singh S, Misra SC, Singh G. Leveraging Additive Manufacturing for Enhanced Supply Chain Resilience and Sustainability: A Strategic Integration Framework. Global Journal of Flexible Systems Management. 2024 Mar 27 [accessed 2024 Apr 11]. https://link.springer.com/10.1007/s40171-024-00389-w. doi:10.1007/s40171-024-00389-w

[14] Weckenborg C, Schumacher P, Thies C, Spengler TS. Flexibility in manufacturing system design: A review of recent approaches from Operations Research. European Journal of Operational Research. 2024;315(2):413–441. doi:10.1016/j.ejor.2023.08.050

[15] Huwa R, Albrecht M, Leinenbach A, Schilp J. Flexibility Dimensions in Cyber-Physical Production Systems. In: 2024 4th International Conference on Electrical, Computer, Communications and Mechatronics Engineering

(ICECCME). 2024. p. 1–6. https://ieeexplore.ieee.org/document/10796674. doi:10.1109/ICECCME62383.2024.10796674

[16] Huwa R, Albrecht M, Leinenbach A, Schilp J. Towards Modelling Flexibility Limits In Cyber-Physical Production Systems. 2025:83–92. doi:10.15488/18858

[17] Wang X, Zhang L, Liu Y, Laili Y. An improved deep reinforcement learning-based scheduling approach for dynamic task scheduling in cloud manufacturing. International Journal of Production Research. 2024;62(11):4014–4030. doi:10.1080/00207543.2023.2253326

[18] Zhang L, Yang C, Yan Y, Hu Y. Distributed Real-Time Scheduling in Cloud Manufacturing by Deep Reinforcement Learning. IEEE Transactions on Industrial Informatics. 2022;18(12):8999–9007. doi:10.1109/TII.2022.3178410

[19] Wang X, Zhang L, Liu Y, Zhao C. Logistics-involved task scheduling in cloud manufacturing with offline deep reinforcement learning. Journal of Industrial Information Integration. 2023;34:100471. doi:10.1016/j.jii.2023.100471

[20] Zhang Q, Li S, Pu R, Zhou P, Chen G, Li K, Lv D. An adaptive robust service composition and optimal selection method for cloud manufacturing based on the enhanced multi-objective artificial hummingbird algorithm. Expert Systems with Applications. 2024;244:122823. doi:10.1016/j.eswa.2023.122823

[21] Jing W, Zhang Y, Chen Y, Zhang H, Huang W. Cloud-edge collaboration composition and scheduling for flexible manufacturing service with a multi-population co-evolutionary algorithm. Robotics and Computer-Integrated Manufacturing. 2024;90:102814. doi:10.1016/j.rcim.2024.102814

[22] Yin C, Li S, Li X. An optimization method of cloud manufacturing service composition based on matching-collaboration degree. The International Journal of Advanced Manufacturing Technology. 2024;131(1):343–353. doi:10.1007/s00170-024-13119-4

[23] Cutting-Decelle AF, Young RIM, Michel JJ, Grangel R, Le Cardinal J, Bourey JP. ISO 15531 MANDATE: A Product-process-resource based Approach for Managing Modularity in Production Management. Concurrent Engineering. 2007;15(2):217–235. doi:10.1177/1063293X07079329

[24] Huwa R, Albrecht M, Leinenbach A, Schilp J. Towards Modelling Flexibility Limits In Cyber-Physical Production Systems. 2025:83–92. doi:10.15488/18858

Biography

Robin Huwa (*1998) holds a position as a research associate at the Chair of Digital Manufacturing at the University of Augsburg since 2024, researching production process virtualization concepts such as cloud manufacturing and equipment-as-a-service.

Marcus Albrecht (*1993) is a member at the Chair of Digital Manufacturing at the University of Augsburg. His research focuses on cyber-physical systems and information sharing in Production with the focus on process optimization.

Andreas Leinenbach (*1991) has been a member of the Chair of Digital Manufacturing at the University of Augsburg since 2023. His research interests include the intelligent production automation.

Johannes Schilp (*1976) is head of the Chair of Digital Manufacturing since 2015. He is also head of the Processing Department at the Fraunhofer Research Institution for Casting, Composite and Processing Technology and president of the Scientific Society of Assembly Handling and Industrial Robotics (WG MHI).

8th Conference on Production Systems and Logistics

Reinforcement Learning for Intercropping Sequences: Transferable Decision Policies for Circular Industrial Production Systems

Stefan Sarawinsky, Kai Schaller, Kajan Kandiah
Institute for Industrial Managament (FIR) at RWTH Aachen University, Aachen, Germany

Abstract

Improving the sustainability of production systems requires decision support methods that can solve the problem of optimizing the combination of resources under ecological, operational, and compatibility constraints [1]. This challenge arises in agriculture and manufacturing, particularly in Circular Industrial Production Systems (CIPS), where the creation of value depends on the coordinated use of materials, processes and resources, while minimizing waste and environmental impact [2]. As production environments become more complex and dynamic, traditional rule-based or static optimization approaches struggle to cope with combinatorial decision spaces and evolving system constraints [3]. To address this issue, the authors use intercropping as an example proxy domain with high complexity that mirrors these combinatorial challenges and compatibility constraints. This paper therefore presents a Deep Q-Network (DQN)-based Reinforcement Learning (RL) approach embedded in a domain-specific simulation environment for planning intercropping sequences. The RL-based Decision Support System (DSS) learns to maximize an aggregated reward comprising agronomic quality indicators and binary user preference feedback. Results from a simulation demonstrate a consistent convergence and structured long-term policy learning in complex combinatorial decision spaces. Beyond agriculture, this framework shows how RL-based sequencing policies can help CIPS facing similar optimization problems involving material combinations, process compatibility, and sustainability trade-offs.

Keywords

Reinforcement Learning; Circular Economy; Decision Support Systems; Sustainable Sequence Optimization

1. Introduction

In agriculture, intercropping is a strategy that enhances soil health, resource efficiency and ecological resilience by combining different plant species within a cultivation system [4]. From a decision-theoretical perspective, intercropping constitutes a highly complex optimization problem [5]. Its success depends on non-linear interactions between plant species, soil properties, seasonal planting windows and operational constraints [5]. A similar level of complexity exists in Circular Industrial Production Systems (CIPS), where planners must identify combinations of materials, processing steps and resource allocations that are compatible with sustainability and cost constraints [6].

In practice, intercropping and planning are not yet widely adopted. One reason for this is the complexity of the planning process based on challenges such as the limited mechanization of work steps and complex interactions between plants [7]. Furthermore, mixed and monoculture planning is still largely driven by expert knowledge, look-up tables, and heuristic rules [8,9]. While these approaches encode valuable domain expertise, they are inherently static and limited in their ability to adapt to new conditions or information, or to incorporate feedback from system operators. Similar limitations are evident in circular industrial

DOI: https://doi.org/10.15488/20994
ISSN: 2701-6277

production planning, where predefined rules and optimization heuristics often fail to capture dynamic interactions and long-term effects of sequential decisions [10].

Consequently, there is a growing need for adaptive, learning-based decision-making methods that can optimize sequences of independent choices over time while incorporating user feedback. Although LLMs excel in semantic reasoning, their inherent stochasticity and challenges in handling delayed rewards mean they are less well-suited than Reinforcement Learning (RL) to reproducible, long-term sequencing problems. RL explicitly optimizes cumulative policies [11,12,13,14]. RL addresses sequential decision-making directly by learning policies that maximize cumulative reward through interaction with the environment. It is particularly suited to problems characterized by multi-step dependencies, uncertainty and delayed feedback [15].

This research aims to evaluate the effectiveness of RL in managing the combinatorial complexity of intercropping as a basis for industrial applications. To this end, we address the following research questions:

RQ1: To what extent can a DQN-based agent internalize and optimize intercropping sequences under heterogeneous agronomic constraints without explicit rule-based programming?

RQ2: How can the structural properties of agricultural intercropping be mapped to CIPS to ensure methodological transferability?

Accordingly, the focus lies on analyzing learning convergence and structural transferability rather than statistical performance benchmarking.

2. Related Work

2.1 Reinforcement Learning in Agriculture

RL has gained increasing attention in the field of agriculture decision support [16,17], particularly in areas such as irrigation management [18,19], fertilization, and crop scheduling [18]. Sutton and Barto established the foundational theoretical framework for RL, setting out the principles of value-based learning, policy optimization, and temporal difference methods that underpin modern RL algorithms [20]. Chen et al. demonstrated the effectiveness of Deep Q-Network (DQN) in irrigation decision-making under uncertain weather conditions, achieving better results than rule-based methods [21]. Gautron et al. provide a thorough overview of RL applications in crop management, emphasizing the suitability for handling delayed rewards and complex interactions while highlighting challenges related to data availability and interpretability [22].

Despite the growing body of work, RL research in agriculture has predominantly focused on single-variable decision problems (e.g. irrigation timing) and monoculture-based systems [18,19]. The optimization of intercropping sequences involving discrete combinatorial decisions and compatibility constraints remains under-explored. Furthermore, there are no studies that explicitly address the transferability of agricultural RL methods for intercrop sequence planning to the context of circular industrial production.

2.2 Decision Support in Circular Industrial Production Systems

In manufacturing, DSS traditionally rely on mathematical programming, heuristics, or simulation-based optimization to address production planning and control problems [23]. However, with the increasing importance of sustainability and circular economy principles, production systems are becoming more complex due to material recyclability, process compatibility, and lifecycle constraints [24,25]. These challenges have motivated research into artificial intelligence (AI)-based decision support, including simulation-driven optimization and machine learning approaches such as RL [26].

Many manufacturing decision-support solutions focus on static optimization or supervised learning methods, which require labeled datasets and struggle to adapt to new operating conditions (e.g. supply chain

disruptions) [26,27]. Sequential decision problems, such as process routing or material sequence optimization, are often addressed through heuristics (e.g. First-Come-First-Served) rather than learning-based approaches [28,29]. This creates a methodological gap that RL is well-positioned to fill, particularly when combined with simulation environments analogous to digital twins.

2.3 Explainability and Human-in-the-Loop Learning

Explainability is a critical requirement for DSS within intercropping practices and circular industrial production. Explainable AI (XAI) methods aim to make model decisions transparent to users. In RL, interpretability can be enhanced through explicit reward decomposition or policy visualization [30]. Compared to end-to-end deep learning, value-based RL approaches such as DQN offer greater opportunities for stepwise explanation of decisions, which is essential for user acceptance [31].

Human-in-the-loop (HITL) learning, where user feedback is incorporated during training, has been shown to improve the alignment of automated decisions with operational preferences [31]. In DQN, reward shaping provides a structured mechanism to embed such feedback into RL systems without compromising learning stability. Despite its potential, this approach remains underutilized in joint agricultural and manufacturing decision support research.

3. Research Methodology

3.1 Problem Abstraction

The intercropping planning problem is formulated as a Markov Decision Process (MDP) defined by the tuple (S, A, R, γ) where S denotes the state space, A the action space, R the reward function, and γ the discount factor. In detail the model abstraction components are:

- **State Space (S):** A state s_t represents a partially filled planting row, encoded by the sequence of crops already placed, the soil type, planting month, and row length.
- **Action Space (A):** An action a_t corresponds to selecting the next crop to be placed at the current position from the available catalog.
- **Implicit Time-Step $(t, t+1)$:** The time index t represents an implicit decision step. Each step corresponds to the placement of one crop in the planting row. $t+1$ denotes the subsequent decision state following this placement.
- **Episode:** An episode ends when the planting row has been completely assigned.

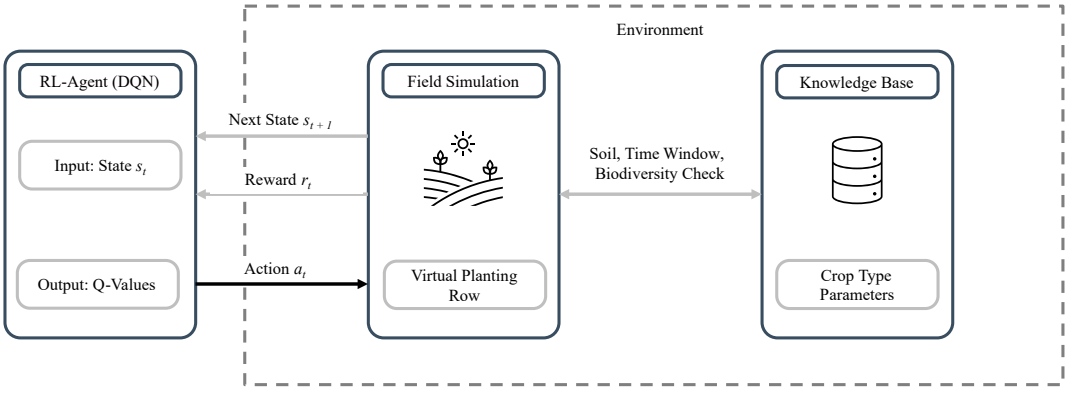


Figure 1: Intercropping approach based on MDP

Figure 1 illustrates the RL loop, in which a DQN agent makes a sequential planting decision based on the current field state s_t. The selected action (e.g. crop index 55: zucchini) is passed to the field simulation in

the environment, where it is validated against an expert knowledge base derived from the agricultural dataset. Within this simulation, specialized filter modules evaluate specific data attributes, such as soil compatibility, seasonal planting windows, and biological synergies, such as nutrient demand. The environment then returns a composite reward r_t, which aggregates agronomic scores for soil and neighborhood effects. It also returns an updated state vector s_{t+1}, which represents the growing planting row. This iterative process enables the agent to learn an optimized sequencing policy (Q-Values) that balances complex operational and ecological constraints effectively.

Simulation Environment

The environment incorporates deterministic expert knowledge through structured data attributes such as soil suitability, planting windows, pH tolerance, nutrient requirements, light demand, pesticide infestation and crop family affiliation. These attributes are used to compute interpretable sub-rewards reflecting agronomic feasibility and diversity.

Reward Structure

The reward function is defined as an additive composition:

$$R_{\text{total}} = R_{\text{agronomic}} + \alpha \cdot R_{\text{user}} \quad (1)$$

where $R_{\text{agronomic}}$ aggregates domain-specific sub-scores (soil compatibility, temporal suitability, nutrient/ light diversity), and R_{user} represents a learned user preference signal derived from binary feedback. The scaling factor α controls the influence of user preferences on the overall optimization objective.

Deep Q-Network Architecture

A DQN is used to approximate the action-value function $Q(s, a)$. The network takes the state vector as input and outputs Q-values for all admissible actions. Training follows standard DQN techniques including experience replay, ε-greedy exploration and temporal-difference learning, as described by Mnih et al. [32].

User Feedback and Reward Shaping

Binary user feedback (acceptable/not acceptable) is collected after each proposed sequence and used to train a lightweight preference model. The output of this model is mapped to a bounded reward adjustment and integrated via online reward shaping. This enables continuous adaptation of the policy to farm specific operational preferences while preserving the underlying agronomic constraints.

3.2 Transferability to Circular Industrial Production Systems

From a methodological perspective, the defined MDP maps directly to CIPS:

Table 1: MDP to CIPS mapping

Intercropping	Circular Industry	Mapping
Crops	Materials	Biological constraints with material compatibility constraints (e.g. in recycling or disassembly)
Planting Rows	Production Sequences	Sequence of crops effects e.g. soil health; the sequence of manufacturing steps affects energy usage and waste generation
Agronomic Constraints	Process Feasibility	pH/Light requirements map to machine capabilities and temperature process windows

As shown in Table 1, crops correspond to material options, planting rows to production sequences, and agronomic constraints to process feasibility in CIPS, regarding compatibility and sustainability requirements. The learned RL policy thus provides a transferable optimization paradigm for sequential decision problems, particularly where combinatorial complexity and sustainability criteria dominate.

4. Experimental Setup

The proposed approach was implemented in Python using a custom OpenAI Gym environment which models a single planting row as a sequential decision problem. The state of the environment is represented as an integer vector of 10 (the maximum length of row), where each position either encodes an empty slot or the index of a plant species from the dataset. The environment models intercropping planning as a finite-horizon decision task, in which a RL agent selects crops incrementally to form a planting sequence, subject to agronomic and operational constraints.

The agronomic knowledge base is stored in a lightweight SQL database containing information on vegetable crops and their attributes. This includes soil suitability scores for three soil types (sand, loam and clay), planting windows (month from/to), pH tolerance ranges, light and nutrient requirements, botanical family and indices for sowing and harvesting techniques. These attributes are used to compute an additive reward comprising (i) soil suitability, (ii) temporal alignment with the planting windows, and (iii) neighborhood effects with respect to the previous plant in the row (light and nutrient complementarity, pH interval overlap, botanical diversity, and process compatibility for sowing and harvesting). Global and local duplicate penalties discourage the repeated use of the same crop.

Each episode represents the planning of a single planting row between 2 and 10 positions long. For each episode, the planting month and soil type are randomly selected. The state is defined as a fixed-length vector encoding the partially filled row, where unassigned positions are padded. The action space is therefore discrete, corresponding to the selection of a crop species from those available at each position.

As mentioned, the reward function is additive and combines multiple agronomic criteria. Duplicate crops within a row incur a penalty. All reward components are logged separately to enable post-hoc explainability.

In addition to expert-based rewards, a lightweight preference model can be activated to incorporate binary user feedback. This model predicts a preference score for individual placement decisions based on contextual features (crop, neighbor, soil and month position), which is mapped to a bounded bonus and added via reward shaping. For the core experiments reported here, the focus is on the base agronomic reward to analyze fundamental learning behavior.

This is how the RL agent and training were formed. The decision policy is learnt using a DQN with two fully connected hidden layers (64 neurons each). Experience replay and a target network with soft updates are employed for stabilization. Training was performed over 15.000 episodes using an ε-greedy exploration strategy that gradually shifted towards exploitation. The total episodic reward is recorded after each episode and serves as the primary performance indicator.

The performance of the RL agent is evaluated based on its behavior as it converges within the complex decision space. The initial training phase (exploration) establishes a foundation for uninformed decision-making. The resulting policy is analyzed in terms of its long-term reward structure and decision-making patterns across different scenarios.

5. Results

The learning behavior of the RL agent based on DQN was evaluated by analyzing the evolution of the total episodic reward over training time.

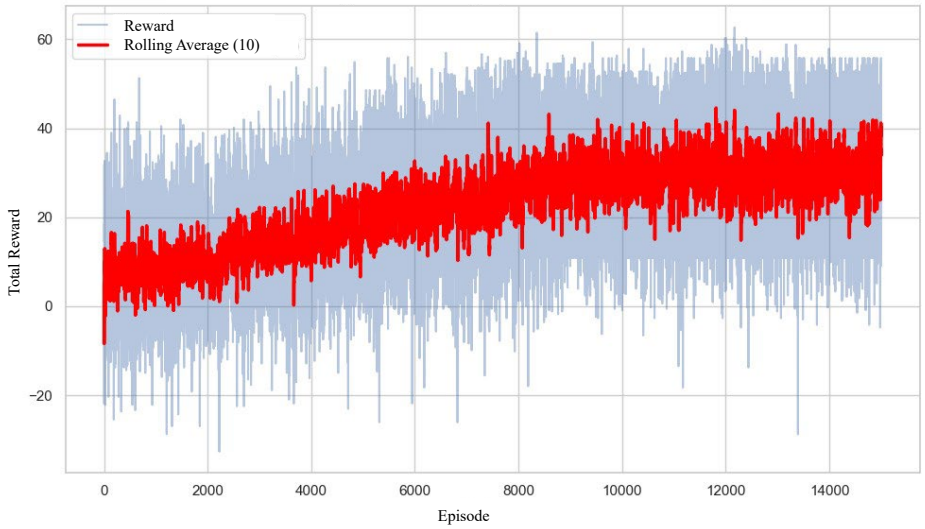


Figure 2: Evolution of episodic rewards during DQN training.

Figure 2 shows the progression of rewards across episodes, displaying both raw episodic rewards and a rolling average over 10 episodes. The rolling mean highlights convergence behavior. Initially, the agent exhibits highly variable rewards, reflecting random exploration and the absence of an informed policy. As training progresses, the rolling average steadily increases, indicating that the agent is learning to exploit the structured agronomic signals embedded in the reward function. After approximately 8.000-10.000 episodes, the reward trend stabilizes, suggesting convergence towards a consistent policy.

Qualitative inspection of sequences generated confirms that the agent learns to avoid planting duplicate crops, to respect planting windows and optimize neighbor compatibility (e.g. by pairing deep-rooting plants with shallow-rooting plants). However, the variance in raw rewards remains relatively high even in later episodes, due to the stochastic sampling of months, soil types, and row lengths. Nevertheless, the monotonic increase in the rolling average demonstrates robust learning progress across heterogeneous scenarios rather than overfitting to a fixed configuration.

The results demonstrate that the DQN agent successfully internalizes multi-criteria agronomic constraints. The steady increase from an initial low reward (representing random exploration) to a stable plateau confirms that the agent learns a policy that demonstrates structured policy learning beyond random exploration. Qualitative analysis confirms that the agent considers long-term neighbor compatibility, which would be impossible for a purely random or short-sighted approach.

6. Discussion

The aim of this study was to determine whether RL with DQN could be an effective method for optimizing complex sequencing problems under heterogeneous constraints. From a methodological perspective, RL differs from greedy heuristics in that it optimizes cumulative reward over a finite planning horizon rather than immediate local effectiveness. This structural property enables the RL agent to consider long-term compatibility effects within the sequence. The following sections discuss the performance of the proposed approach, its methodological transferability to CIPS and the limitations of the current study.

6.1 Performance and Contributions

The experimental results indicate that the RL approach successfully internalizes multi-criteria constraints without relying on explicit rule-based programming. The convergence of the cumulative reward (as shown in Figure 2) suggests that the RL agent learns to balance conflicting objectives, such as maximizing diversity while ensuring soil compatibility, by exploring the structure of the reward function. In contrast to static rule-

based heuristics that optimize based on predefined local criteria, the RL agent optimizes cumulative reward over the full planning horizon.

A key strength of the proposed approach lies in its interpretability. By decomposing the reward into individual agronomic components and logging them for every placement decision, the system provides transparent explanations for its recommendations. This is particularly important in practitioner-facing DSS, where trust and domain acceptance are critical.

A critical requirement for DSS in both agriculture and industry is explainability. The proposed framework addresses this by utilizing a decomposed reward structure. Since every decision score is an aggregation of transparent sub-rewards, the system can provide post-hoc explanations for its recommendations. This transparency is essential for building trust with domain experts, who are often skeptical of black-box optimization methods.

6.2 Transfer to Circular Industrial Production Systems

A central contribution of this work is the validation of intercropping systems as a high-complexity proxy domain for CIPS. Although the semantic differs, the underlying mathematical structure of the decision problem is isomorphic between the two domains. Optimizing a sequence of discrete entities under strict compatibility and sustainability constraints applies for both approaches.

6.2.1 Structural Isomorphism

The methodological transferability rests on the following mapping between the intercropping agriculture approach and CIPS. Crops correspond to material options, planting rows to production sequences, and agronomic constraints to process feasibility (see Table 1). Sustainability objectives, such as soil fertility preservation in agriculture, are mirrored by circularity metrics in industry, including material reuse, energy consumption and waste reduction.

Just as crops have biological attributes (pH, nutrients), industrial components possess material attributes (alloy type, recyclability, toxicity) that dictate compatibility [33]. The order of crops affects soil health and pest resilience, analogously, the sequence of manufacturing or disassembly steps determines energy consumption, tool changes, and material purity [6]. Constraints such as planting windows map to temporal production restrictions (e.g. curing times, shift schedules) or machine capabilities [10].

6.2.2 Sustainability Trade-offs

Both domains face the challenge of multi-objective optimization where economic efficiency conflicts with ecological goals. In intercropping, the RL agent balances yield against soil regeneration. In CIPS, planners must balance throughput against circularity metrics, such as maximizing the reuse of components or minimizing hazardous waste [2]. The results of this paper demonstrate that RL is particularly suited for this problem because it enables sustainability as a cumulative reward signal rather than a hard constraint, allowing the system to find flexible trade-offs that rigid rule-based system often misses.

6.2.3 Human-in-the-Loop with Reward Shaping

The integration of user feedback highlights a crucial mechanism for industrial transferability. Production planners, like farmers, possess tacit knowledge [31]. The proposed reward shaping mechanism allows the RL agent to adapt these latent preferences over time. This capability is necessary for the deployment of CIPS, where soft factors like worker experience or varying material quality often deviate from the theoretical model.

6.3 Limitations

While the results are promising, several limitations must be acknowledged to contextualize the findings:

- **Simulation and Reality (L1):** The evaluation relies on a simulation environment with deterministic biological interactions. Real-world biological and industrial systems exhibit a higher degree of stochasticity (e.g. weather anomalies or machine breakdowns) that were not fully modeled.
- **Planning Horizon (L2):** The current optimization focuses on single-row sequences within a single planting season. It does not yet account for multi-period crop rotations or the complex interdependencies between parallel production lines, which would exponentially increase the state space.
- **Feedback Data (L3):** The user preference module was tested with limited binary feedback data. In a real-world deployment, the agent would lack sufficient feedback in the start phase, which would result in not aligned decisions. This remains a challenge that requires more robust initialization strategies, such as imitation learning from historical logs.

Despite these limitations, the study confirms that RL-based sequencing policies provide a robust alternative to static heuristics. The successful application in the intercropping domain serves as a proof-of-concept for the broader class of sustainable sequencing problems, offering a methodological blueprint for intelligent decision support in circular economy.

7. Conclusion & Outlook

This paper investigated the application of RL to optimize intercropping sequences, using the agricultural context as a methodological proxy domain for CIPS to address complex operational constraints. A DQN was embedded in a domain-specific simulation environment and trained using an interpretable, multi-criteria reward function. The experimental results show that the agent learns stable sequencing policies that internalize non-trivial compatibility relationships between crops and exhibit structured long-term optimization behavior.

Beyond its agricultural contribution, the work demonstrates that intercropping constitutes a valuable high-complexity proxy domain for sustainability-oriented decision support. The methodological principles underlying the proposed RL framework by sequential policy learning, reward decomposition and HITL adaptation. Therefore, it's directly transferable to circular industrial production systems facing similar optimization challenges.

The paper highlights several strengths of RL-based decision support: the ability to balance conflicting objectives over time, adaptability to changing constraints, and inherent support for explainable decision-making through structured reward signals. At the same time, limitations related to simulation fidelity (L1), limited feedback data (L3) and simplified planning horizons (L2) underline the need for further research.

Future work will extend the framework along three main directions. First, the environment will be expanded to more complex multi-row and multi-period planning scenarios to capture longer-term interactions and rotations. Second, RL will be systematically compared with non-deterministic LLM-based decision approaches, investigating also hybrid architectures that combine symbolic reasoning with sequential optimization. Third, the transfer potential will be empirically validated through an industrial case study in a circular production context, such as remanufacturing or disassembly sequencing.

To summarize, this work provides both a concrete RL-based solution for intercropping planning and a transferable methodological contribution for adaptive, sustainability-oriented decision support. As production systems increasingly adopt circular economy principles, RL offers a promising foundation for next-generation intelligent planning and control systems.

Acknowledgements

Funded as DLR project no. 01MN23012B by the Federal Ministry of Research, Technology and Space (BMFTR) due to a decision of the German Parliament.

References

[1] Ghadimi, P., Wang, C., Lim, M.K., 2019. Sustainable supply chain modeling and analysis: Past debate, present problems and future challenges. Resources, Conservation and Recycling 140, 72–84.

[2] Geissdoerfer, M., Savaget, P., Bocken, N.M., Hultink, E.J., 2017. The Circular Economy – A new sustainability paradigm? Journal of Cleaner Production 143, 757–768.

[3] Helo, P., Hao, Y., 2022. Artificial intelligence in operations management and supply chain management: an exploratory case study. Production Planning & Control 33 (16), 1573–1590

[4] Malézieux, E., Crozat, Y., Dupraz, C., Laurans, M., Makowski, D., Ozier-Lafontaine, H., Rapidel, B., Tourdonnet, S., Valantin-Morison, M., 2009. Mixing plant species in cropping systems: concepts, tools and models. A review. Agron. Sustain. Dev. 29 (1), 43–62

[5] Dury, J., Schaller, N., Garcia, F., Reynaud, A., Bergez, J.E., 2012. Models to support cropping plan and crop rotation decisions. A review. Agron. Sustain. Dev. 32 (2), 567–580

[6] Lieder, M., Rashid, A., 2016. Towards circular economy implementation: a comprehensive review in context of manufacturing industry. Journal of Cleaner Production 115, 36–51.

[7] Raseduzzaman, M., Jensen, E.S., 2017. Does intercropping enhance yield stability in arable crop production? A meta-analysis. European Journal of Agronomy 91, 25–33.

[8] Memmah, M.-M., Lescourret, F., Yao, X., Lavigne, C., 2015. Metaheuristics for agricultural land use optimization. A review. Agron. Sustain. Dev. 35 (3), 975–998.

[9] Saranya, S., Amudha, T., 2016. Crop planning optimization research — A detailed investigation, in: 2016 IEEE International Conference on Advances in Computer Applications (ICACA). 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, India. 24.10.2016 - 24.10.2016. IEEE, pp. 202–208.

[10] Dolgui, A., Ivanov, D., Sokolov, B., 2018. Ripple effect in the supply chain: an analysis and recent literature. International Journal of Production Research 56 (1-2), 414–430.

[11] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., Zhou, D., 2022. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models.

[12] Valmeekam, K., Marquez, M., Olmo, A., Sreedharan, S., Kambhampati, S., 2022. PlanBench: An Extensible Benchmark for Evaluating Large Language Models on Planning and Reasoning about Change.

[13] Herrera-Poyatos, D., Peláez-González, C., Zuheros, C., Herrera-Poyatos, A., Tejedor, V., Herrera, F., Montes, R., 2025. An overview of model uncertainty and variability in LLM-based sentiment analysis: challenges, mitigation strategies, and the role of explainability. Frontiers in artificial intelligence 8, 1609097.

[14] Wang, H., Zhu, Z., Shi, F., 2025. Distribution Prompting: Understanding the Expressivity of Language Models Through the Next-Token Distributions They Can Produce.

[15] Kaelbling, L.P., Littman, M.L., Moore, A.W., 1996. Reinforcement Learning: A Survey. jair 4, 237–285.

[16] Bharman, P., Ahmad Saad, S., Khan, S., Jahan, I., Ray, M., Biswas, M., 2022. Deep Learning in Agriculture: A Review. AJRCoS, 28–47.

[17] Chen, D., Huang, Y., 2025. Integrating reinforcement learning and large language models for crop production process management optimization and control through a new knowledge-based deep learning paradigm. Computers and Electronics in Agriculture 232, 110028.

[18] Saikai, Y., Peake, A., Chenu, K., 2023. Deep reinforcement learning for irrigation scheduling using high-dimensional sensor feedback.

[19] Yu, J., Qu, Q., Peng, S., Wei, X., Li, Y., Sun, C., 2025. Deep learning for intelligent irrigation decision-making: A review. Agricultural Water Management 320, 109836.

[20] Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. IEEE Trans. Neural Netw. 9 (5), 1054.

[21] Chen, M., Cui, Y., Wang, X., Xie, H., Liu, F., Luo, T., Zheng, S., Luo, Y., 2021. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. Agricultural Water Management 250, 106838.

[22] Gautron, R., Maillard, O.-A., Preux, P., Corbeels, M., Sabbadin, R., 2022. Reinforcement learning for crop management support: Review, prospects and challenges. Computers and Electronics in Agriculture 200, 107182.

[23] Pinedo, M.L., 2022. Scheduling. Springer International Publishing, Cham.

[24] Khan, I.S., Ahmad, M.O., Majava, J., 2021. Industry 4.0 and sustainable development: A systematic mapping of triple bottom line, Circular Economy and Sustainable Business Models perspectives. Journal of Cleaner Production 297, 126655.

[25] Machado, C.G., Winroth, M.P., Da Ribeiro Silva, E.H.D., 2020. Sustainable manufacturing in Industry 4.0: an emerging research agenda. International Journal of Production Research 58 (5), 1462–1484.

[26] Li, C., Zheng, P., Yin, Y., Wang, B., Wang, L., 2023. Deep reinforcement learning in smart manufacturing: A review and prospects. CIRP Journal of Manufacturing Science and Technology 40, 75–101.

[27] Usuga Cadavid, J.P., Lamouri, S., Grabot, B., Pellerin, R., Fortin, A., 2020. Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. J Intell Manuf 31 (6), 1531–1558.

[28] Hubbs, C.D., Li, C., Sahinidis, N.V., Grossmann, I.E., Wassick, J.M., 2020. A deep reinforcement learning approach for chemical production scheduling. Computers & Chemical Engineering 141, 106982.

[29] Waschneck, B., Reichstaller, A., Belzner, L., Altenmüller, T., Bauernhansl, T., Knapp, A., Kyek, A., 2018. Optimization of global production scheduling with deep reinforcement learning. Procedia CIRP 72, 1264–1269.

[30] Ruggeri, F., Russo, A., Inam, R., Johansson, K.H., 2025. Explainable Reinforcement Learning via Temporal Policy Decomposition.

[31] Christiano, P., Leike, J., Brown, T.B., Martic, M., Legg, S., Amodei, D., 2017. Deep reinforcement learning from human preferences.

[32] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I.,

King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518 (7540), 529–533.

[33] Vongbunyong, S., Chen, W.H., 2015. Disassembly Automation. Springer International Publishing, Cham.

Biography

Stefan Sarawinsky

Stefan Sarawinsky is a researcher at the Institute for Industrial Management (FIR) at RWTH Aachen University. His research focuses on machine learning, agent-based AI, and the design and management of scalable IT architectures. Since joining FIR in 2024, he has been actively involved in research on reinforcement learning in the agronomic domain, particularly within the SAAT project. He holds a master's degree in computer science.

Kai Schaller

Kai Schaller is a research assistant at the Institute for Industrial Management (FIR) at RWTH Aachen University. His research activities primarily focus on artificial intelligence, with particular emphasis on machine learning methods. Since joining FIR in 2025, he has been engaged in research on reinforcement learning applications in the agriculture domain. He holds a bachelor's degree in industrial engineering from RWTH Aachen University.

Kajan Kandiah

Kajan Kandiah has been a researcher at the Institute for Industrial Management (FIR) at RWTH Aachen University since 2020. In his current role within the Information Management department, he is part of the technical group IT-complexity management. His expertise includes IT architecture management, smart grid technologies and transformation strategies.

8th Conference on Production Systems and Logistics

A Digital Framework For Structured Meeting Preparation In Kaizen Projects

Kevin Zhao, Yuxi Wang, Joachim Metternich
PTW/ TU Darmstadt, Darmstadt, Germany

Abstract

Unstructured meetings are a major source of efficiency-related losses in organizations. To address this, a role-based, multi-agent meeting support framework powered by large language models (LLMs) is introduced to enhance structured preparation for cross-functional collaboration. In this framework, each agent simulates a domain-specific role (e.g., logistics, assembly, machining), contributing specialized knowledge to generate targeted agendas and prepare contextual materials aligned with standard operating procedures (SOPs). The approach was evaluated in a learning factory environment using a simulated Kaizen case study, demonstrating its ability to generate goal-driven agendas. Initial results show that well-structured agenda creation is possible, and that the framework can support goal-oriented, cross-functional meetings by generating actionable agendas and providing relevant, explainable information for stakeholder selection. By improving agenda relevance and reducing coordination overhead, this framework provides a scalable method to integrate intelligent meeting support into continuous manufacturing improvement processes, offering a novel application of LLMs to support cross-functional continuous improvement initiatives in manufacturing.

Keywords

Lean Manufacturing; Continuous Improvement; Large Language Models; Agenda Automation; Cross-functional collaboration

1. Introduction

Of the many inefficiencies that exist in organizations, meetings are especially prominent due to their significant time consumption and clear potential for enhancement. They represent a significant cost factor, especially for managerial staff [1]. For instance, in one case study, 300 managers collectively spend an estimated 60,000 hours annually in meetings. At an average personnel cost of €45 per hour, this results in a total annual meeting cost of approximately €2.7 million [1]. In addition to the financial impact, meeting quality influences employee job satisfaction [2], which has been shown to correlate with customer satisfaction [3]. Beyond these human and financial implications, effective meetings are also challenged by organizational structures. Manufacturing organizations typically consist of production and support departments, each with distinct objectives and specialized domain knowledge. However, this division can hinder effective collaboration across departments due to limited mutual understanding and a heightened risk of miscommunication [4]. A lack of coordination and integration of different departments in the internal supply chain can lead to lower performance and numerous non-value-adding activities, including delays due to waiting, time spent searching for information, and participation in non-goal-oriented meetings [5,6,4,7]. These inefficiencies contribute to increased transactional costs that are rarely captured in commercial accounting systems [6].

DOI: https://doi.org/10.15488/20995
ISSN: 2701-6277

The analysis of meetings to enhance their effectiveness has been the focus of extensive research for several decades. A consistent finding across these studies is that thorough preparation significantly enhances both the effectiveness and inclusivity of meetings [8–10]. In particular, the use of a clearly defined agenda and the distribution of pre-meeting materials are identified as critical factors that enable participants to prepare adequately and contribute meaningfully to the discussion [8–10]. Given these implications, there is a clear need for a standardized and systematic approach for meeting preparation in manufacturing environments. This need is particularly critical in Kaizen improvement projects, in which different departments need to work together to realize a shared objective [11]. However, despite extensive research on Kaizen methods and plan-do-check-act (PDCA) processes, existing literature provides almost no guidance on how Kaizen project meetings themselves should be systematically prepared, highlighting a clear gap in structured approaches to cross-functional meeting preparation. To address this gap, we propose a framework that supports organizers by automatically generating structured agendas and relevant pre-meeting information based on a given meeting topic. The prototype compiles contextual background knowledge and identifies the required expertise, enabling a relevant participant list. By automating these preparatory steps, the system reduces coordination effort and allows teams to focus on value-adding analysis and decision-making.

2. State of the art

2.1 Kaizen

Kaizen is a philosophy that translates to “change for the better” and is widely recognized as continuous improvement [12]. From this core philosophy, numerous concepts, and systems, like the PDCA cycle, have been developed to help implement Kaizen within organizations, with their effective implementation depending on teamwork, communication, and employee morale, among other factors. The PDCA cycle is a systematic approach used to drive continuous improvement within the framework of Kaizen. It begins with planning, which involves identifying and defining a specific objective for improvement. The do phase focuses on implementing the planned intervention. During the check phase, results are analyzed by comparing actual outcomes against the planned objectives to determine whether the intervention has achieved measurable improvement. Finally, in the act phase, successful plans are standardized to prevent recurrence of the original issue, and the insights gained are used to guide future cycles of improvement [11].

There have been several reports of successful Kaizen projects, that follow the PDCA cycle. For instance, Chen et al. [13] describe an application of a Kaizen-based methodology in the design of cellular manufacturing systems. In their study, improvement projects were initiated by identifying existing operational problems, followed by a structured sequence that included team formation, data collection, idea generation, and implementation of solutions. Specifically, the objective was to reduce production costs of meat tenderizers while concurrently enhancing product quality. Upon project completion, results were presented to management, thereby not only validating the project’s organizational impact but also reinforcing a culture of continuous improvement throughout the enterprise.

Despite these successes, the reported cases primarily reflect individual Kaizen initiatives implemented within specific organizational contexts. For example, Aktar et al. [14] emphasize that their findings cannot be universally generalized across industries or sectors, as their research was limited to a single company producing a specific type of product. This limitation underscores the highly context-specific nature of Kaizen initiatives, which often poses challenges to operational standardization and broader scalability. Furthermore, not all Kaizen projects achieve their intended outcomes. As highlighted by Bader et al. [15], common reasons for failure include inadequate communication, insufficient allocation of resources, and broader organizational barriers.

2.2 Meeting preparation

One critical yet often overlooked component of effective collaboration is the proper planning and execution of meetings. Meetings occupy a considerable portion of employees' time and constitute a significant organizational cost. To maximize their effectiveness, careful planning is essential [16]. There are several factors that help with the planning of a meeting. When the information can be communicated quickly, clearly, and effectively via email, this option should be chosen over scheduling a meeting [16]. However, if a meeting is necessary, it is important to invite only those individuals whose participation is directly relevant. Engagement tends to be higher when participants have a clear and meaningful role and are able to contribute to the agenda items [16,17]. Establishing a clear and concise agenda is essential for setting a productive tone at the outset of any meeting. Even a basic agenda adds to the success of meetings [18].

2.3 Multi-Agent Large Language Model System

The definition of the term *agent* has been the subject of ongoing scholarly debate. In this work, the definition proposed by Woolridge [19] is adopted, in which an agent is described as "a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives."

A distinction is typically made between single-agent and multi-agent environments. In a single agent environment, an agent operates independently to achieve its goal, for instance, solving a crossword puzzle without external interaction. In contrast, multi-agent environments involve multiple agents whose actions may influence one another. A classic example is a game of chess, which inherently involves at least two agents interacting strategically [20].

The concept of agents can be extended to the domain of LLMs, where each LLM can be considered as an individual agent capable of autonomously performing a wide range of tasks. When multiple LLMs interact and collaborate, they collectively form a multi-agent system. Such multi-agent LLM systems are particularly compelling, as recent research has demonstrated their potential in addressing complex problem-solving scenarios [21].

2.4 Retrieval-Augmented generation

LLMs do not possess inherent access to proprietary or domain-specific knowledge, such as internal company data. To enable LLMs to generate contextually relevant and goal-oriented responses grounded in such information, external data must be integrated into their processing pipeline [22]. According to Gao et al. [22] two main approaches are used to provide LLMs with access to external knowledge. The first is fine-tuning, which involves training the model on a selected dataset that reflects the target domain. This process updates the model's internal parameters so the model can generate responses based on the newly learned domain-specific knowledge. Although this approach yields significant results, it requires extensive computational resources and is associated with high costs. The second approach is retrieval-augmented generation (RAG), which does not modify the model itself. Instead, it retrieves relevant information from an external knowledge base and incorporates it into the model's input when generating a response.

3. Methodology

3.1 System Overview

An overview of the system architecture is presented in Figure 1. Conceptually, the system is structured into three core components. The first, User Interaction, allows users to engage with the multi-agent system through natural language prompts, requiring minimal effort from the user beyond providing meeting information. The second component, Multi-Agent Discussion, comprises multiple LLM agents, each representing a different department and having access to that department's knowledge domain. This allows the system to determine which departments are relevant to a meeting, facilitate targeted collaboration, and

identify participants who may not need to be invited. The final component, Pre-Meeting Material, is triggered once the agent discussion concludes. In this stage, the system follows a structured conversation protocol to produce a discussion summary. The agenda is then generated from this summary, ensuring that only relevant information is included. Both the summary and agenda are delivered back to the user. This architecture ensures a seamless interaction cycle, empowering users to effortlessly generate essential preparatory materials for meetings that can be shared with relevant stakeholders.

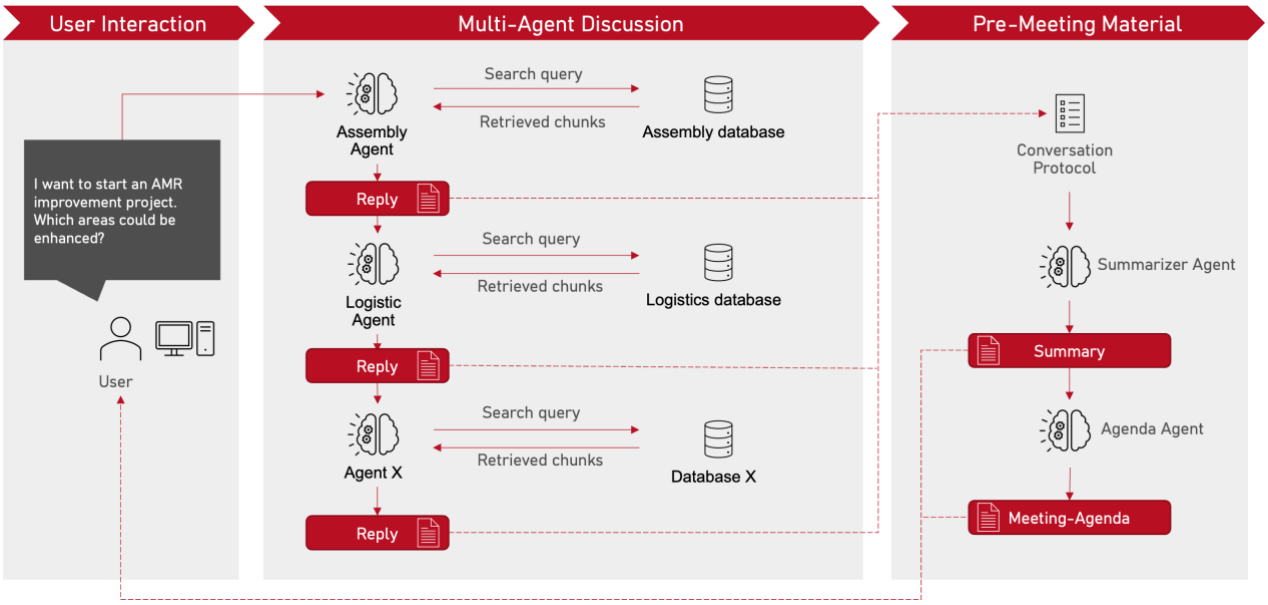


Figure 1: System overview illustrating the complete workflow from User Interaction to the generation and distribution of meeting materials, shown using an example meeting topic on improving Autonomous Mobile Robots (AMR)

The underlying idea of this architecture is to support Kaizen-oriented workflows by standardizing meeting preparation across the PDCA cycle, extracting relevant information to distribute to the team, and helping identify which departments should be involved in the meeting. In doing so, the system is designed to reduce administrative effort and to minimize the risk of involving stakeholders who cannot meaningfully contribute, thereby allowing teams to allocate their time and resources more effectively toward continuous improvement activities rather than coordination tasks.

3.1.1 User Interaction

The user interacts with the system through a graphical interface in which a prompt can be entered. This prompt should contain information about the meeting topic the organizer intends to schedule. At this stage, the user may also provide optional contextual details that are not covered by the existing knowledge domain but may enrich the subsequent agent discussion. Once the prompt is submitted, the user's involvement is complete. The system then processes the input autonomously and returns the generated reports. For transparency, the user can additionally review the multi-agent conversation history to understand how specific information in the output was produced.

3.1.2 Multi-Agent Discussion

A central contribution of the prototype is its design of orchestrated LLM agents, each representing a distinct organizational department, collaboratively engaged in information retrieval and inter-agent communication. These agents operate within their respective domains of knowledge to simulate intra-organizational consultation and information retrieval. This multi-agent architecture ensures that agenda items and preparatory materials are comprehensive, accurate, and traceable to their originating departments, thereby increasing transparency and accountability in the preparation process. The meeting facilitator then can decide on basis of the output of the system who to involve in the meeting.

In our system, the conversation logic follows a round-robin approach, where each agent is called sequentially. The process begins when the user submits a prompt, which is first received by the initial agent. This agent automatically performs an information retrieval process within its departmental knowledge base to identify relevant data. Based on the retrieved information, the agent then generates a response. If no relevant information is found, the agent produces a response indicating its inability to contribute, which is recorded in the protocol. This ensures that the final report reflects which departments have relevant knowledge, aiding the meeting facilitator in deciding whether certain stakeholders should be involved. After the first agent responds, the user input along with the conversation history is passed to the next agent in sequence. Each agent follows the same process and formulates its own search query based on the preceding conversation to identify relevant information. This cycle continues for each defined agent until all agents have responded, after which a new round begins. Otherwise, once the predefined number of rounds is completed, the agent discussion concludes.

3.1.3 Pre-Meeting Material

After the discussion concludes, the entire conversation protocol is forwarded to a summarizer agent. This agent produces a condensed summary that highlights the specific contributions made by each department within the multi-agent system. This ensures that all relevant information is captured and presented in a way that is easier for users to understand. Based on this summary, the subsequent agenda agent generates a meeting agenda. The agenda design follows the planning checklist proposed by Leblanc and Nosik [16] which outlines a form to be completed by the meeting leader both during planning and after the meeting. The resulting system-generated agenda should clearly state the purpose of the meeting and specify the contributions of each participating department.

3.2 Evaluation Strategy

The evaluation framework is designed to assess both the performance of the multi-agent system and the degree to which the generated summary and agenda adhere to the log. To this end, we employ the metrics outlined in Table 1, which jointly measure output quality and agent-level interaction quality. Our focus in this work is to evaluate the generated summaries, agendas, and the underlying multi-agent discussion dynamics. By analyzing each aspect independently, we can pinpoint specific areas for improvement and establish a foundation for systematic benchmarking, including comparisons against alternative architectures such as single-agent baselines.

Table 1: Metrics regarding the evaluation strategy with description of the purpose of the measurement

Evaluation Aspect	Evaluation Type	Metric	What it Measures	Source
Summary Quality	Coverage	G-Eval	Does the summary include the important content from the log?	[23]
Summary Quality	Faithfulness	AlignScore	Does the summary avoid hallucinating or distorting the log?	[24]
Agenda vs Summary	Agenda Coverage	G-Eval	Does the agenda capture the main topics from the summary?	[23]
Agenda Quality	Structure Quality	DiscoScore	Does the agenda maintain logical coherence and sentence ordering?	[25]
System Quality	Agent Consistency	BERTScore	Do the agents' contributions exhibit semantic agreement?	[26]

While our system employs RAG for information access, we do not include RAG-specific metrics in the evaluation. Our focus is on overall summary and agenda quality as well as multi-agent coordination, and detailed optimization of RAG-based retrieval is beyond the scope of this paper.

4. Experimental Evaluation

4.1 Setup

For the User Interface layer, we employ Open WebUI [27], an open-source platform that supports the integration of locally hosted LLMs via Ollama. Open WebUI provides a user-facing environment for submitting prompts and interacting with the underlying system. We extend its functionality by adding custom pipelines that enable seamless interaction with our multi-agent architecture. To orchestrate the multi-agent discussion, we build on the AutoGen framework [28], with which agents can be created. When a user issues a prompt, the interface displays how an agent gets called including invoked search queries and generated replies, along with their retrieved sources and chunks.

For all agent computations, we employ the llama3.3:70b model provided through Ollama, running on an NVIDIA RTX Pro 6000 Blackwell GPU equipped with 96 GB VRAM.

4.2 Case Study: Kaizen Project Meeting Preparation

To evaluate the effectiveness of the proposed system, we conducted a case study within a learning factory designed to replicate an industrial value stream. The factory produces pneumatic cylinders through a sequence of turning, milling, cleaning, assembly, and intra-logistics processes. This environment enables realistic cross-functional interactions while remaining controllable for research purposes. For the case study, we designed a 60-minute meeting scenario focused on the continuous improvement of existing standard operating procedures (SOPs) in a Kaizen project context. The scenario takes place in the plan phase of the PDCA cycle, during which participants are expected to analyze current practices, identify deviations and inefficiencies, and define a roadmap for targeted improvements. To guide the agents, we formulated the following meeting prompt: „*The goal of the meeting is to analyze the current standards, identify deviations and inefficiencies, and define a prioritized roadmap with targeted improvement measures*.".

The scenario involved three domain-specific agents: a Logistics Agent responsible for material flow and internal logistics coordination; an Assembly Agent focusing on the worker assistance system and the efficiency of assembly operations; and a Plant Agent overseeing the end-to-end value stream and identifying systemic weaknesses based on inputs from the other agents. The selection of these three agents reflects the structure of the available SOP documentation in the learning factory, which can be grouped into the areas of logistics, assembly, and plant-level process coordination. The agents' knowledge base consists of SOPs from a learning factory, which we cluster by department to reflect organizational structure. As they are written in German language, the conversation protocol as well as summary and agenda will also be in German language to avoid cross-lingual inconsistencies and ensure that the model processes the input in its original linguistic form.

4.3 Results

Figure 2 presents an excerpt from the agent discussion as well as the corresponding User Interface. In this example, the Logistics Agent first performs a RAG-based retrieval to identify relevant information from its knowledge base, then displays the retrieved sources and generates its response accordingly.

Table 2 shows the agenda created by the multi-agent system in response to the user's prompt. Based on the input, the system formulated the meeting title as "Efficiency Improvement in the Logistics Department" and defined the objective as: "The goal of this meeting is to discuss and implement concrete measures to improve efficiency in the logistics department through the introduction of automation, quality control, and continuous

improvement processes." The table presents the agenda items as generated by the system, including timing, goals, and responsible participants. The note at the end, also produced by the system, states: "The meeting is intended to be completed within 60 minutes and to provide a clear guideline for future efficiency improvements in the logistics department." The table reflects an English translation of the original agenda created by the system.

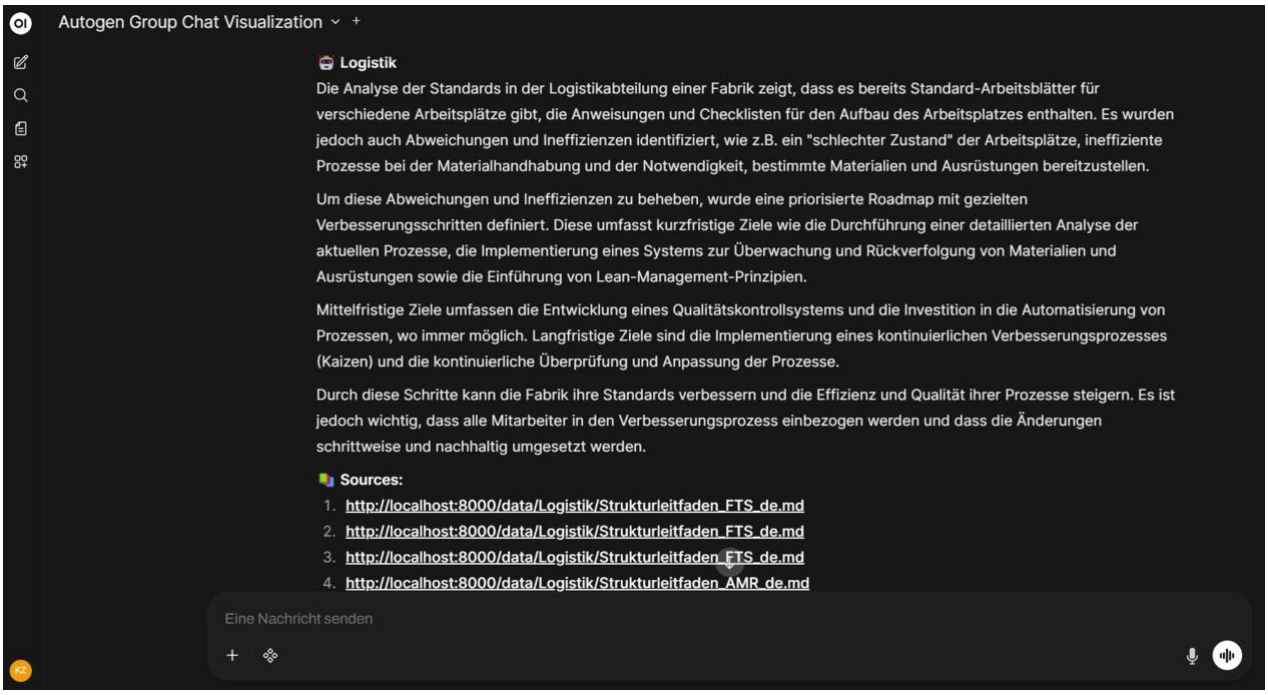


Figure 2: Screenshot of the OpenWebUI User Interface displaying one instance of an agent response alongside the sources it used

Table 2: Agenda created by the multi-agent system, including responsible participants (English translation)

Time	Agenda Item	Goal	Responsible
5 min	Introduction and Objectives	Introduce the topic and clarify the goals of the meeting	Moderator
15 min	Presentation of Current Challenges and Opportunities	Provide an overview of current challenges and identify opportunities for efficiency improvements	Logistics Department
20 min	Presentation of Automation Options (FTS/AMR) and Quality Control	Explain the potential of driverless transport systems (FTS), autonomous mobile robots (AMR), and quality control systems for increasing efficiency	Engineering / IT Department
10 min	Discussion on Implementation	Discuss necessary steps for implementing proposed measures and the role of each team member	Moderator
10 min	Next Steps & Conclusion	Define concrete tasks, responsibilities, and deadlines for implementing the agreed measures	Moderator

Because the evaluation metrics used in this study were originally developed for English, we translate the generated German summaries and agendas into English using Neural Machine Translation (OPUS-MT [29]), prior to evaluation. German versions of G-Eval, AlignScore, and DiscoScore are not currently available, so translation serves as the only consistent and established basis for comparison across all metrics. Although a

German BERT-based metric exists, applying it exclusively to the original German outputs while using other metrics on translated content would introduce methodological inconsistency. While translation may introduce minor distortions, this approach ensures comparability with prior work and preserves methodological integrity.

Table 3: Resulting metrics of the Case Study of Kaizen Planning in a learning factory

Evaluation Aspect	Evaluation Type	Metric	Range	Score
Summary Quality/ Agenda Quality	Coverage	G-Eval	[1, 10]	5.8
Summary Quality	Faithfulness	AlignScore	[0, 1]	0.4846
Agenda vs Summary	Coverage	G-Eval	[1, 10]	8.5
Agenda Quality	Structure Quality	DiscoScore (DS_Focus_NN)	[0, ∞)	1.1038
Agenda Quality	Structure Quality	DiscoScore (DS_SENT_NN)	[-1, 1]	0.9361
System Quality	Agent Consistency	BERTScore (mean)	[0, 1]	0.8161

The results of the case study are summarized in Table 3. The summary achieved a G-Eval Score of 5.8, indicating that it captures the main content of the multi-agent discussion, although it lacks the actionable detail requested in the user query. In contrast, the agenda performs more strongly, showing clearer structure and better alignment with the key topics.

The AlignScore (0.4846) indicates partial faithfulness, revealing some deviations from the discussion log, though the metric does not specify whether these stem from omissions or paraphrasing. In terms of structural quality, the agenda exhibits strong sentence-level semantic connectivity, as reflected by its DS_SENT_NN score (0.9361), while the DS_Focus_NN score (1.1038) suggests only modest divergence in the distribution of key concepts relative to the summary.

Finally, Agent Consistency, measured using mean BERTScore (0.8161), shows that the agents' responses are generally consistent and build upon each other. It is important to note that excessively high values could indicate redundancy, where agents reiterate the same points rather than providing diverse perspectives.

5. Discussion and Outlook

The results demonstrate that the proposed framework can generate agendas that are logically structured based on a provided summary, offering a solid baseline for meeting facilitators to design targeted meetings with clear agendas and actionable points. Additionally, the system provides relevant documents along with explainable outputs, allowing users to identify the source department and select appropriate stakeholders for the meeting. The framework is designed to allow the integration of additional agents representing further departments. The focus of the generated agenda on logistics-related topics reflects the distribution of relevant information within the SOP corpus used for retrieval. This behavior is intentional, as the agenda aligns with the departments and issues that the system identifies as most strongly related to the meeting objective.

As a first prototype, the system is not optimized for performance, but it demonstrates the feasibility and practical relevance of applying multi-agent LLMs to Kaizen meeting preparation, an area previously unaddressed in the literature. The main limitation lies in the faithfulness of the generated summaries, which suggests potential for improving inter-agent coordination, experimenting with alternative orchestration logics, or refining prompt engineering strategies.

Future work will focus on validating the approach through additional case studies, integrating practitioner feedback, and benchmarking against single-agent and baseline systems. These steps will further clarify the system's impact on meeting quality and continuous improvement processes in real industrial settings.

Acknowledgements

The authors gratefully acknowledge the financial support of the research project "PrePAIR" (project number 13IK037I) funded by the European Union and supported by the German Federal Ministry for Economic Affairs and Energy (BMWE) on the basis of a decision by the German Bundestag.

References

[1] Rausch, A., 2009. Steigerung der Meeting Performance als Managementaufgabe. Controlling & Management 53 (6), 376–383.

[2] Rogelberg, S.G., Allen, J.A., Shanock, L., Scott, C., Shuffler, M., 2010. Employee satisfaction with meetings: A contemporary facet of job satisfaction. Human Resource Management 49 (2), 149–172.

[3] Chi, C.G., Gursoy, D., 2009. Employee satisfaction, customer satisfaction, and financial performance: An empirical examination. International Journal of Hospitality Management 28 (2), 245–253.

[4] Institut für angewandte Arbeitswissensch (Ed.), 2016. 5S als Basis des kontinuierlichen Verbesserungsprozesses. Springer Berlin Heidelberg, Berlin, Heidelberg.

[5] Bertagnolli, F., 2020. Lean Management: Einführung und Vertiefung in die japanische Management-Philosophie. Springer Fachmedien Wiesbaden, Wiesbaden.

[6] Ebber, R., Lennings, F., 2020. Lean-Administration – das verschenkte Potenzial! Betriebspraxis und Arbeitsforschung 240, 31–42.

[7] Narasimhan, R., Das, A., 2001. The impact of purchasing integration and practices on manufacturing performance. Journal of Operations Management 19 (5), 593–609.

[8] Cutler, R., Hosseinkashi, Y., Pool, J., Filipi, S., Aichner, R., Tu, Y., Gehrke, J., 2021. Meeting Effectiveness and Inclusiveness in Remote Collaboration. Proceedings of the ACM on Human-Computer Interaction 5 (CSCW1), 1–29.

[9] Leach, D.J., Rogelberg, S.G., Warr, P.B., Burnfield, J.L., 2009. Perceived Meeting Effectiveness: The Role of Design Characteristics. Journal of Business and Psychology 24 (1), 65–76.

[10] Rogelberg, S.G., Shanock, L.R., Scott, C.W., 2012. Wasted Time and Money in Meetings: Increasing Return on Investment. Small Group Research 43 (2), 236–245.

[11] Imai, M., 2012. Gemba Kaizen : a commonsense approach to a continuous improvement strategy, 2nd ed. ed. McGraw Hill Global education, New York.

[12] Palmer, V.S. Inventory management Kaizen, in: Proceedings 2nd International Workshop on Engineering Management for Applied Technology. EMAT 2001. IEEE Comput. Soc, Austin, TX, USA, pp. 55–56.

[13] Chen, J.C., Dugger, J.C., Hammer, R., 2001. A Kaizen Based Approach for Cellular Manufacturing System Design: A Case Study. Journal of Technology Studies 27 (2), 19–27.

[14] Aktar Demirtas, E., Gultekin, O.S., Uskup, C., 2023. A case study for surgical mask production during the COVID-19 pandemic: continuous improvement with Kaizen and 5S applications. International Journal of Lean Six Sigma 14 (3), 679–703.

[15] Bader, M., Antony, J., Jayaraman, R., Swarnakar, V., Goonetilleke, R.S., Maalouf, M., Garza-Reyes, J.A., Linderman, K., 2024. Why do process improvement projects fail in organizations? A review and future research agenda. International Journal of Lean Six Sigma 15 (3), 664–690.

[16] LeBlanc, L.A., Nosik, M.R., 2019. Planning and Leading Effective Meetings. Behavior Analysis in Practice 12 (3), 696–708.

[17] Ludwig, T.D., Frazier, C.B., 2012. Employee Engagement and Organizational Behavior Management. Journal of Organizational Behavior Management 32 (1), 75–82.

[18] Niederman, F., Beise, C.M., Beranek, P.M. Issues and Concerns about Computer-Supported Meetings: The Facilitator's Perspective. MIS Quarterly 20 (1), 1.

[19] Wooldridge, M., 2002. An Introduction to MultiAgent Systems. John Wiley & Sons, Chichester.

[20] Russell, S.J., Norvig, P., 2016. Artificial intelligence: a modern approach, Third edition, Global edition ed. Pearson, Boston Columbus Indianapolis.

[21] Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N.V., Wiest, O., Zhang, X., 2024. Large Language Model based Multi-Agents: A Survey of Progress and Challenges. arXiv. doi:10.48550/arXiv.2402.01680. Accessed 9 May 2025.

[22] Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., Wang, H., 2024. Retrieval-Augmented Generation for Large Language Models: A Survey. arXiv. doi:10.48550/arXiv.2312.10997. Accessed 23 July 2025.

[23] Liu, Y., Iter, D., Xu, Y., Wang, S., Xu, R., Zhu, C., 2023. G-Eval: NLG Evaluation using GPT-4 with Better Human Alignment, 11 pp. https://arxiv.org/pdf/2303.16634v3. Accessed 2 December 2025.

[24] Zha, Y., Yang, Y., Li, R., Hu, Z., 2023. AlignScore: Evaluating Factual Consistency with a Unified Alignment Function, 19 pp. https://arxiv.org/pdf/2305.16739v1. Accessed 2 December 2025.

[25] Zhao, W., Strube, M., Eger, S. DiscoScore: Evaluating Text Generation with BERT and Discourse Coherence, 3865–3883.

[26] Zhang, T., Kishore, V., Wu, F., Weinberger, K.Q., Artzi, Y., 2019. BERTScore: Evaluating Text Generation with BERT, 43 pp. https://arxiv.org/pdf/1904.09675v3.

[27] Baek, J., Hussain, A., Liu, D., Vincent, N., Kim, L.H., 2025. Open WebUI: An Open, Extensible, and Usable Interface for AI Interaction. arXiv. doi:10.48550/arXiv.2510.02546. Accessed 9 October 2025.

[28] Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A.H., White, R.W., Burger, D., Wang, C. AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation.

[29] Tiedemann, J., Thottingal, S., 2020. OPUS-MT - Building open translation services for the World, in: Proceedings of the 22nd Annual Conference of the European Association for Machine Translation. European Association for Machine Translation, Lisboa, Portugal, pp. 479–480.

Biography

Kevin Zhao M.Sc. (*1997) is research associate and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt since 2024. His research centers around data science and AI in production, especially for digitalized shopfloor processes.

Yuxi Wang, M.Sc. (*1995) is research associate and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt since 2021. Her research interests include the further development and introduction of natural language processing to digital shop floor management and problem-solving systems**.**

Prof. Dr.-Ing. Joachim Metternich (*1968) is head of the PTW-Institute at the Technical University of Darmstadt. He is the spokesman of the medium-sized businesses digital competence center Darmstadt; he has been president of the International Association of Learning Factories (IALF) from 2015 to 2021 and he is scientific partner of the Conference on Production Systems and Logistics (CPSL).

8th Conference on Production Systems and Logistics

Structuring the Shift: Designing and Applying a Smart PSS Transformation Framework with the TriProGenix Approach

Jonas Wirth[1,3,4], Dennis Unic[2,3], Thomas Bauernhansl[1,4]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12, 70569 Stuttgart, Germany*
[2]*Aalen University - Technology, Business and Health, Beethovenstr. 1, 73430 Aalen*
[3] *TRUMPF, Johann-Maus-Straße 2, 71254 Ditzingen, Germany*
[4] *Institute of Industrial Manufacturing and Management (IFF), University of Stuttgart*

Abstract

In today's industrial landscape, companies are experiencing mounting pressure to enhance their product offerings through digital innovation and services. A significant number of companies are transitioning from conventional product-centric business models to Smart Product-Service Systems (SPSS), which seamlessly integrate physical products with data-driven services, thereby delivering enhanced value. However, the process of transitioning to a smart PSS is accompanied by several significant challenges. In many cases, substantial organizational and procedural transformation is a prerequisite. Companies encounter difficulties in aligning diverse disciplines and managing increased system complexity, and they must adapt internal structures to support new smart service offerings. To address the challenges previously mentioned, this paper proposes a transformation framework that systematically guides the design and implementation of smart PSS. The framework integrates product, software, and service development aspects within a unified business model context, ensuring technical solutions and service components co-evolve in alignment with the company's value proposition. The proposed framework is anchored in the TriProGenix approach. This structured methodology has been developed to facilitate the transformation of industrial offerings into SPSS across a huge range of application domains. The framework was derived from a comprehensive analysis of extant development methodologies and practical applications. This paper presents one of the four industrial case studies that were used to operationalize and validate the TriProGenix approach, thereby helping to refine it further. The following discourse will be furnished with insights derived from an industrial case study. The application of the framework is demonstrated in a manufacturing context, illustrating how it facilitates the integration of a data-driven service component into existing operational processes. Adopting the proposed framework enables practitioners to methodically navigate the digital servitization process, thereby effectively bridging the gap between traditional product development and smart service integration. This, in turn, accelerates the transformation towards sustainable smart PSS offerings.

Keywords

Smart Product-Service Systems (SPSS); Digital Servitization; Transformation Framework; Business Model Innovation; TriProGenix-Model

1. Introduction/Motivation

Mechanical and plant engineering is a key industrial sector with an annual turnover of around €269 billion and almost one million employees [1,2]. At the same time, the industry faces profound structural change driven by saturated markets, increasing international competition, and a growing shortage of skilled workers

DOI: https://doi.org/10.15488/20996
ISSN: 2701-6277

[3–7]. The traditional business model of one-off machine sales is reaching maturity, with stagnating revenues and decreasing differentiation [6]. Industry 4.0 technologies such as IoT, artificial intelligence, cloud platforms, and big-data analytics enable new forms of automation, real-time monitoring, and data-driven optimization [8–11]. They provide the basis for hybrid and service-oriented business models that extend far beyond pure product sales [12–14]. In this context, manufacturers increasingly complement physical products with digital services to realize additional customer value and new revenue streams, leading to a shift towards data-driven, usage-based business models [15–18]. Service-oriented, data-based concepts such Product-Service Systems (SPSS) are therefore gaining importance but are often still in early stages, particularly in mechanical and plant engineering. Many companies struggle to translate technological innovation into scalable and economically viable offerings due to a lack of systematic transformation frameworks [5,19].

Against this background, the objective of this paper is to develop and empirically validate a structured approach for transforming a product-centric, AI-based assistance system into a data-driven Smart Product-Service System (Smart PSS). This leads to the guiding research question: *"How can the TriProGenix approach be leveraged to develop and validate a transformation-oriented framework (ProGenLab-X) for converting a product-centric AI assistance system into a data-driven Smart PSS?"* Within the authors' broader TriProGenix research program on data-driven and service-oriented business models in mechanical and plant engineering, ProGenLab-X represents the Smart-PSS-specific instantiation of the TriProGenix process logic and serves as a phase-based reference model for transforming product-centric offerings into scalable SPSS.

2. Fundamentals

2.1 Servitization as a basis for developing new business models

Servitization describes the strategic shift from product-centric to service-oriented value creation. Vandermerwe & Rada argue that integrating services into traditional manufacturing enables sustainable differentiation [20,18]. As services become central elements of business models rather than add-ons [21], companies must adapt structures and routines that often still reflect product logic and therefore create organizational and cultural barriers [22–24]. Nevertheless, servitization is seen as a key transformation path to build long-term competitive advantages [16,14,18].

2.2 Business models under the influence of digital transformation

Digital transformation introduces new value creation logics enabled by IoT, AI, cloud computing, and big data. These technologies support continuous optimization, data-driven services, and automated decision-making, forming the basis for digital and hybrid business models [9,25,26]. The convergence of digitalization and servitization (digital servitization) requires organizations to adapt IT infrastructures, processes, and competencies to integrate data-driven services at scale [27,28,24]. Data becomes a central resource, enabling insights, performance improvements, and usage- or results-based revenue models [7,29,14].

2.3 Digital Ecosystems

With increasing interconnectivity, companies operate within digital ecosystems: dynamic networks of partners, technologies, and platforms that jointly deliver value [30–33]. Ecosystems foster resource sharing, enable integration of data-driven services, and accelerate scaling [34,35,8]. As coordination across technical and organizational interfaces becomes essential, digital ecosystems form the strategic architecture for implementing smart, data-based business models [35,8,25]. For mechanical engineering firms, embedding products and services within such ecosystems is increasingly a prerequisite for competitive service offerings [16,36,37].

2.4 PSS in mechanical and plant engineering

Product-Service Systems (PSS) integrate physical products with accompanying services and form a core pillar of service-oriented business models [37–39]. Smart PSS extend this logic by adding data-driven functionalities such as monitoring, optimization, and AI-based assistance, creating continuous value during operation [40,9]. Enabled by digital technologies, remote services, condition monitoring, predictive maintenance or usage-based billing, they open new opportunities for efficiency, monetization, and customer retention [8,25,41]. However, implementing PSS in mechanical and plant engineering remains challenging due to organizational, skill-related, and infrastructural requirements [22,42,19]. Although relevance is widely recognized, many firms are still in early stages of PSS adoption [16,5,19], highlighting the need for structured transformation frameworks that integrate technical, organizational, and business model perspectives [5,22,43]. The TriProGenix approach addresses precisely this gap [19].

In conclusion, it has been demonstrated that the transition towards data-driven SPSS requires more than technological capabilities alone. Successful transformation in this context is dependent on the integration of three factors: service-oriented value creation, logic based on digital business models, and collaboration within ecosystems. Concurrently, mechanical engineering companies must surmount considerable organizational, cultural, and structural impediments if these objectives are to be realized. These findings highlight the need for a structured transformation framework that systematically integrates customer value, data logic, and organizational feasibility and thus provides holistic support for change.

3. Research Methodology

This paper uses the design science research (DSR) methodology according to Hevner et al. [44] to develop a transformation-oriented framework for SPSS. DSR offers a structured approach to designing artefact-based solutions and combines scientific rigor with practical relevance. Accordingly, the research process is divided into three core elements: knowledge base, environment, and design process [44,45]. The knowledge base is based on a systematic literature review in accordance with the procedure described by vom Brocke et al. [46]. To this end, the paper searches the Web of Science, Scopus, IEEE Xplore, and ScienceDirect databases using an extended search string covering data-driven business models, servitization, digital ecosystems, and SPSS. This is followed by duplicate removal, title and abstract screening, and full-text review. In addition, forward and backward searches are performed to identify further relevant publications. In addition, unstructured searches are included to incorporate practice-oriented sources outside of traditional databases. The entire search and selection process is shown in Figure 3 in Appendix 1. The systematic search yielded 455 initial results; following the removal of duplicates, 359 publications were retained. The screening of titles and abstracts resulted in the identification of 101 full texts, of which 47 were deemed relevant. Subsequent forward and backward searches expanded the corpus to 102 publications. Combined with 20 unsystematic sources, this resulted in a final set of 63 publications forming the theoretical foundation of the framework. Moreover, the study incorporates validation insights from ten semi-structured expert interviews, ensuring practical relevance and methodological triangulation.

Furthermore, the development of the artefact was accompanied by a three-stage application-oriented validation procedure, consisting of (1) self-application to assess internal consistency, (2) evaluation by methodologically trained junior experts to test transparency and comprehensibility, and (3) expert validation within the industrial context to ensure practical relevance and applicability. This multi-layer validation constitutes an integral component of the DSR process, thereby substantiating the robustness of the resulting framework.

The industrial environment was examined in cooperation with TRUMPF using an AI-based machine assistance system as the application scenario. The AI component is representative of the digital functionality of the case, yet it does not constitute the central research object of this paper. Instead, it serves as an

illustrative industrial example to demonstrate the transformation towards a Smart PSS offering. Technological, organizational, and business-model requirements derived from ten expert interviews and internal process analyses were integrated into the framework design via the DSR relevance cycle. The iterative development followed the DSR design cycle, covering problem formulation, goal definition, artefact design, demonstration, and evaluation [44] and was refined through expert workshops [47], consistent with methodological approaches in comparable data-driven business model research [44,19]. Combining a systematic literature review, real industrial context, and iterative refinement ensured a scientifically rigorous yet practically applicable transformation framework for smart PSS development, including the adjustment and optimization of cutting parameters.

4. TriProGenix-Framework

The TriProGenix Framework (TPG) is a practice-oriented process model designed for the development and implementation of data-driven and service-oriented business models in mechanical and plant engineering. It provides a structuring navigation system that supports companies in identifying requirements, developing solutions, business models, and anchoring them organizationally [19,48–50]. Figure 1 illustrates the complete framework, integrating its three core dimensions: development levels, process logic, and method categories. As illustrated in Figure 5, the procedural application path is further delineated.

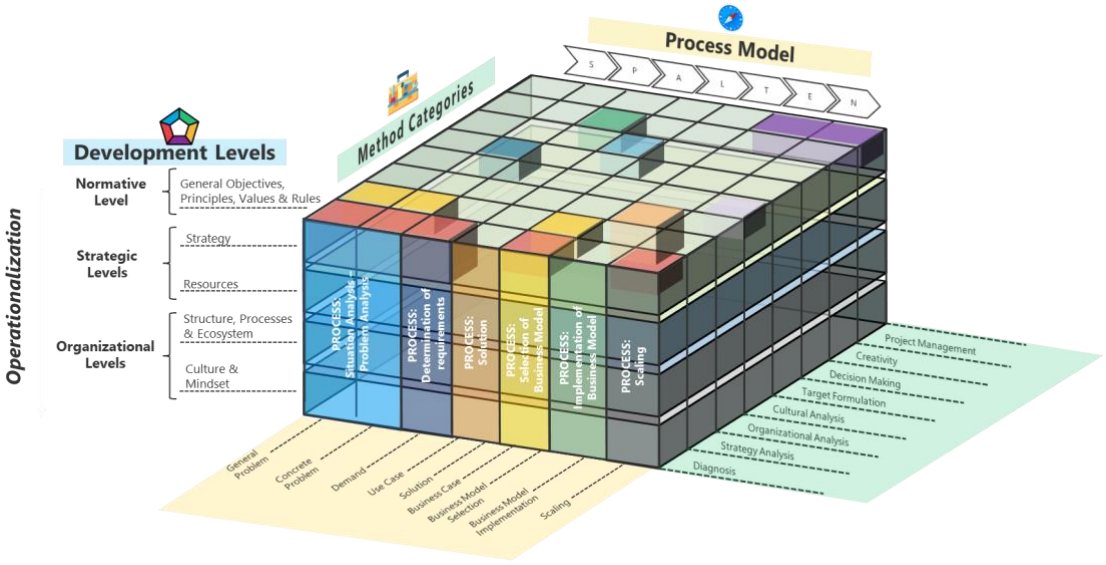


Figure 1: Overview of the TriProGenix method [19].

The first dimension comprises vertical development levels that structure the key areas of action for business model design. At the normative level, companies define target visions, values, and guiding principles. The strategic level addresses service-oriented strategies, suitable business model logics, and resource commitments, while the organizational level focuses on structures, processes, roles, ecosystem relationships, and a supporting culture [19]. The second dimension is a six-step problem-solving process, situation analysis, requirements definition, solution development, business model selection, implementation, and scaling, based on generic approaches such as the SPALTEN system. This process provides a clear path from initial analysis to organizational implementation and scaling of data-driven, service-oriented solutions [51]. The third dimension is a methodological toolkit comprising eight categories that bundle instruments for strategy work, diagnosis, organizational and cultural analysis, decision-making, creativity, and project management to systematically support the process steps [19].

TPG integrates these three dimensions into a holistic transformation framework that links development levels, process logic, and methods, offering a structured approach for converting data-driven products into service-oriented business models. In this paper, TPG serves not only as a conceptual reference but as the

methodological foundation: its process logic and instruments guide the analysis of the initial situation, the structuring of requirements, and the development of the Smart PSS transformation approach. Consequently, the ProGenLab-X framework is derived from the underlying TPG process architecture and operationalizes its principles for the Smart PSS context.

5. Related Research

This paper first consolidates the theoretical and practical requirements for SPSS and uses them as a reference point to assess existing business-model approaches and Frameworks. Theoretical requirements can be structured into content-related, methodological, and formal dimensions [52,53]. Content-related requirements encompass the conception of a (S)PSS as an integrated lifecycle-oriented service bundle [27,54,25,29,38,18], the development of a holistic and continuously delivered value proposition [17,38,39], the establishment of long-term customer relationships across the usage lifecycle [37,17,38], and the creation of measurable performance improvements in the industrial production environment [55,28,25,56]. Methodological requirements emphasize organizational agility [4,37,24]. Formal requirements comprise modularity and flexibility [31,37,33,57], risk-adequate contracting [43,39], monetization logic, and integration into data-driven ecosystems [31,8,36,33]. Practical insights from the industrial case confirm these requirements but additionally highlight validated target markets, demonstrable customer value, subscription-ready IT/ERP infrastructures, and digitalized usage- and billing-data pipelines as essential prerequisites.

Based on this consolidated requirement profile, the paper compares ten established business-model approaches (Figure 4), including frameworks, process models, and canvas-based methods. These approaches were selected because they either play a central role in scientific discourse or are explicitly tailored to hybrid, digital, or service-oriented business models in mechanical and plant engineering. Appendix 1: Framework comparison (Figure 4) summarizes the analyzed approaches and maps their coverage across the consolidated requirement dimensions. The comparative analysis shows substantial variation in both scope and depth.

Generic business-model approaches, such as the Tukker Typology (1) [39], the Business Model Canvas (5) [58] and the Business Model Navigator (9) [16], provide conceptual clarity and help structure PSS variants or business-model building blocks, yet remain too generic to capture data-integration requirements, organizational readiness, or Smart-PSS monetization mechanisms.

PSS-specific approaches, such as the Smart PSS Framework by Kohtamäki et al. [59], the PSS Lifecycle Framework [60] and the PSS Layer Method [61], emphasize the lifecycle-oriented integration of product and service components. However, they largely omit subscription logic, ERP/IT integration, usage-based billing mechanisms, and organizational transformation requirements.

Specialized approaches, including the Business Model Innovation Process (BMIP) [62], the MTC Monetization Canvas [63], the Data-Driven Innovation Framework [11], and the Equipment-as-a-Service Canvas [5], address isolated aspects, innovation dynamics, revenue mechanics, data-driven value creation, or EaaS-specific requirements. Yet, they remain thematically narrow and do not provide a holistic transformation logic that links strategic orientation, organizational design, data infrastructure, and business-model implementation. As demonstrated in Figure 4, none of the analyzed approaches fully covers the complete requirement set derived from theory and practice. The application of comparative mapping has revealed that the extant models address individual requirement clusters, albeit in a fragmented manner. Generic business-model approaches provide a conceptual framework; however, they are deficient in terms of the depth required for IT integration, ecosystem orchestration, and the operationalization of data-driven services. PSS-specific approaches demonstrate stronger alignment between product and service lifecycles, but exhibit clear gaps in monetization logic, organizational implementation capabilities, and the integration of data-enabled service components. Conversely, monetization- and EaaS-oriented approaches provide detailed insights into pricing structures and revenue mechanics. However, these approaches remain

disconnected from broader servitization logic, organizational change processes, and Smart-PSS transformation requirements.

The analysis demonstrates that none of the extant approaches achieve full coverage across all requirement categories, with most scoring only partially or moderately, frequently below 50%. Despite the extensive documentation of the challenges of digital servitisation and Smart PSS implementation in previous research, comparatively little work provides integrated approaches that support companies in systematically overcoming these challenges in practice. This underscores a notable gap in extant research: despite the existence of numerous frameworks addressing discrete components of Smart PSS development, a comprehensive, domain-specific transformation logic remains elusive, one that would facilitate unification of strategic positioning, organizational prerequisites, technical infrastructure, value-proposition design, and monetization mechanisms. The ProGenLab-X framework, which is the focus of this paper, addresses this gap by providing a structured, application-oriented, and industry-compatible model for transforming product-centric offerings into scalable, data-driven SPSS. Embedded within the TPG approach, it operationalizes the generic transformation logic of TPG for the Smart-PSS domain and integrates strategic positioning, organizational prerequisites, technical infrastructure, value proposition design, and monetization mechanisms into a coherent, phase-based model.

6. Framework

The methodological point of departure for the artefact developed in this study is the solution space articulated by the TPG approach. Within the TPG *Solution* pathway, the subprocesses of solution development, result integration, and model selection constitute the conceptual substrate from which the ProGenLab-X framework emerges as a domain-specific instantiation for data-driven SPSS. The framework operationalizes the theoretical premises and practice-derived constraints consolidated in the knowledge base, translating them into an application-oriented process model that guides the migration from product-centric offerings to service-oriented, data-driven business models.

ProGenLab-X exhibits a dual structure comprising a *target path* and a *requirements path*. The target path specifies objectives at the normative, strategic, and business-model levels, whereas the requirements path enumerates the technical, organizational, and market prerequisites necessary to realize each phase. This duality enables transparent gap diagnostics and provides a clear orientation mechanism for consistent transformation aligned with the TPG solution logic.

The framework comprises five iterative phases:

1. **Opportunity identification.** Define relevant markets, use cases, and value levels; delineate initial technological and organizational boundary conditions.
2. **Value-proposition definition.** Formulate a measurable, data-based solution promise; derive the capability requirements necessary for its realization.
3. **Offering and revenue design.** Translate the value proposition into a modular bundle of product, software, and service components; evaluate pricing and revenue mechanisms (e.g., subscriptions, usage-based models).
4. **Infrastructure preparation.** Address operational and technical enablement, including IT/ERP integration, data pipelines, telemetry architecture, service processes, and governance structures.
5. **Market-fit assurance.** Execute pricing, enable sales, establish onboarding routines and customer education, and implement mechanisms for long-term retention.

Figure 2 *(ProGenLab-X SPSS framework)* provides a visual representation of these phases and their alignment with the corresponding requirement categories. Iterative loops across analysis, design, and

implementation are explicitly supported, consistent with the relevance, design, and rigor cycles characteristic of the design-science research paradigm (detailed version see Appendix 4).

Within the scope of this research, the TPG approach is leveraged in two complementary capacities. First, ProGenLab-X serves as a phase-based reference model that instantiates the TPG *Solution* path for SPSS, systematizing the typical steps, artefacts, and decision logics along an opportunity-to-market trajectory. Second, the same model operates as a practical design guide for developing a concrete SPSS business model for an AI-based laser-cutting assistance system, encompassing service bundling, pricing logic, target-segment definition, and initial infrastructure and rollout requirements.

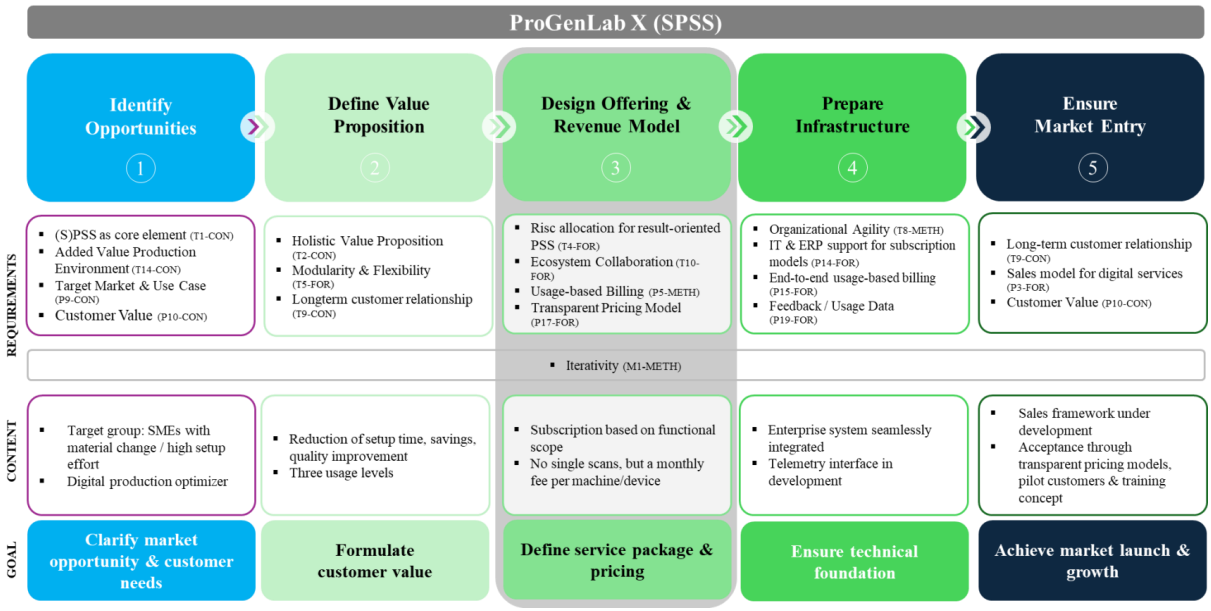


Figure 2: ProGenLab X (SPSS) Framework

In this dual role, ProGenLab X both operationalizes the TriProGenix logic within the Smart PSS domain and demonstrates its applicability in an industrial setting. Consequently, the framework provides an integrated, practice-oriented, and scientifically grounded approach to the development of data-driven SPSS. Its distinguishing characteristics include a modular architecture, the explicit separation of objectives and prerequisites, and its consistent anchoring in TPG principles, features that enable adaptation across diverse industrial contexts and support organizations in the systematic planning, implementation, and scaling of service-oriented business models.

7. Application and Validation

To demonstrate the practical applicability of the ProGenLab-X framework, it was applied to an industrial use case involving an AI-based laser-cutting assistance system that had previously been positioned mainly as a product feature. Using the framework, relevant target markets, prioritized use cases, and central pain points were derived and translated into a clear, value-based proposition. In particular, step three of the framework resulted in the development of a structured service package logic. This logic link identified requirements to sustainable, implementable bundles. Based on this, alternative service bundles, such as usage-based models, subscription packages, and hybrid variants, were designed and linked to required technical and organizational prerequisites. The application shows that the framework provides a consistent structure to connect data-driven service functions with business-model logic and operational implementation requirements, enabling the derivation of a scalable Smart PSS architecture. The validation process was conducted in accordance with a three-stage procedure. Initially, the model was subjected to a self-application process, with the objective of ensuring internal coherence, completeness, and robustness. Secondly, the

evaluation of transparency, comprehensibility, and reproducibility was conducted by junior experts who had received methodological training and who were independent of the domain knowledge of the researchers. The respondents' feedback confirmed the clarity and modularity of the process and led to minor improvements in guidance and visualization. Thirdly, domain experts from the industrial context confirmed the framework's relevance, practical applicability, and compatibility with existing organizational and ecosystem structures. It was emphasized that the model supports the systematic development of monetizable service bundles. Furthermore, it was highlighted that the derived package structure facilitates clear, operationally actionable offerings. Besides the necessity for enhanced stakeholder engagement, notably customers, service units, and information technology, to bolster operational relevance was emphasized. Overall, the application and validation demonstrate that ProGenLab-X provides a methodologically consistent and practically actionable approach for transforming product-centric solutions into data-driven Smart-PSS offerings. It generates concrete, implementable business model options, including structured package logics, and offers a structured pathway that answers the research question regarding the applicability and relevance of the developed artefact in an industrial context.

8. Results and Outlook

The application of the ProGenLab-X framework demonstrates that a structured, TPG-based approach provides substantial value for transforming product-centric industrial solutions into data-driven SPSS. The case study confirms that the framework systematically addresses key transformation challenges, such as unclear value propositions, missing monetization logic, lack of service-oriented organizational structures, and translates these into coherent and actionable business-model designs. By integrating strategic goal definition, requirement derivation, and a consistent process logic across the TPG development levels, the framework enables the development of concrete service bundles and operationally feasible design options applicable to real industrial environments. The three-stage validation confirms the methodological consistency and practical applicability of the framework. Users, junior experts, and domain experts highlighted its clarity, modularity, and reproducibility, while also pointing to improvement potential regarding more operational guidance and a stronger involvement of additional stakeholders such as customers, service, and IT. These insights underline that Smart-PSS transformation requires the interplay of three success factors: data-driven value creation, service-oriented business-model logic, and organizational feasibility. The case further shows that the TPG approach provides an effective methodological foundation for structuring such transformations. The ProGenLab-X framework combines strategic, organizational, and cultural requirements into an integrated transformation logic that supports both customer-oriented value creation and internal implementability. Hence, the guiding research question, how TPG can be used to design and validate a transformation-oriented Smart-PSS model, can be answered positively.

Several limitations remain. The compressed validation format restricted the depth of assessment, as not all process steps could be fully explored, and the empirical basis reflects only a limited set of internal perspectives. Customer, service, and IT viewpoints should therefore be incorporated more comprehensively in future evaluations. Furthermore, while strategically relevant, the strategic layer of the framework has not yet been fully operationalized in the demonstrator. Looking forward, applying the framework in additional industrial domains will be crucial to assess its generalizability and derive cross-sector refinements. Strengthening the integration of strategic elements, such as enterprise-architecture alignment and market positioning, offers considerable potential to enhance managerial impact. Iterative development through Minimum-Viable-Product cycles appears particularly promising to continuously validate and expand data-driven value propositions. Overall, the results demonstrate that the TriProGenix approach constitutes a powerful and transferable structure for guiding companies toward sustainable, data-driven, and service-oriented business models.

Appendix

Appendix 1: Data collection process comparison

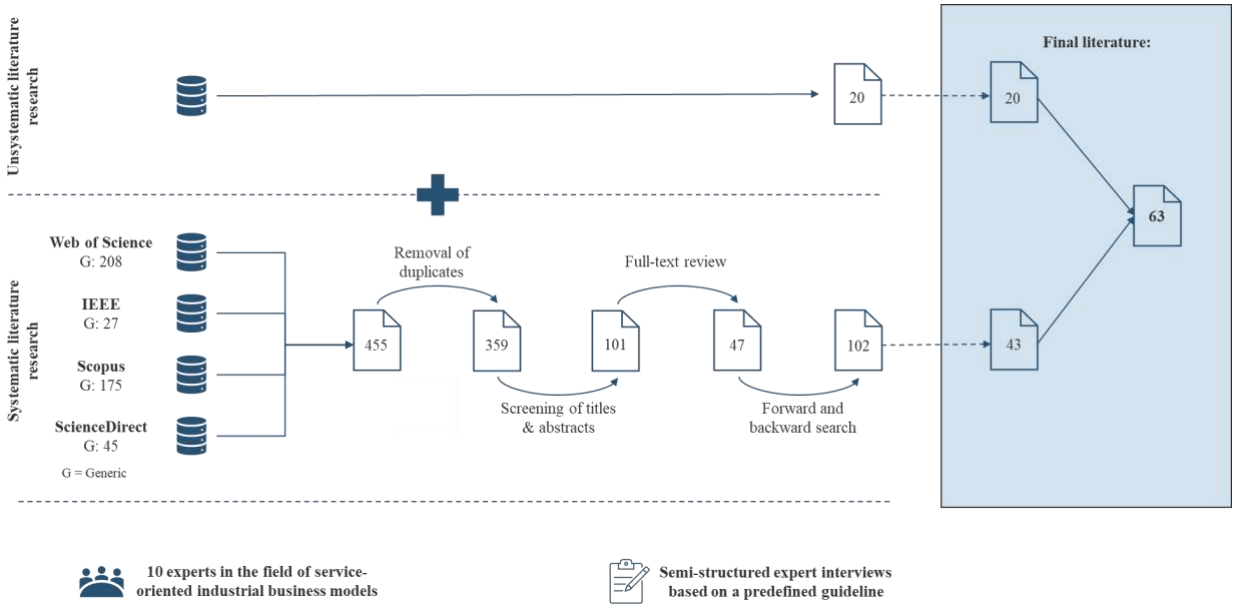


Figure 3: Structured literature research, according to VOM BROCKE et al. [36]

Appendix 2: Framework comparison

	Approaches	Content Requirements						Formal Requirements								Methodical Requirements		
		(S)PSS as core element	Holistic value proposition	Long-term customer relationsship	Added value production environment	Targetmarket & Use Case	Customer Value	Risc allocation for result-oriented PSS	Modularity & flexibility	Ecosystem Collaboration	Sales model for digital services	IT & ERP support for subscription models	End-to-end usage-based billing	Transparent pricing model	Feedback / usage data	Usage-based billing	Organizational agility	**Overall Fulfilment**
1	Tukker-Typologie Tukker (2004)	◕	◕	◔	○	○	◔	○	◔	○	○	○	○	○	○	○	○	○
2	Smart PSS Framework Kohtamäki et al. (2020)	◕	◕	◕	◔	◔	◕	◔	◕	◕	◔	◔	◔	◑	◕	◔	◕	◑
3	PSS-Lifecycle Framework Aurich et al. (2006)	◕	◕	◑	◔	◑	◕	◔	◑	◔	○	○	○	○	○	◔	◔	◔
4	PSS-Layer Method Müller et al. (2009)	◕	◕	◔	◔	○	◕	◔	◕	◔	○	○	○	○	○	○	◔	○
5	Business Model Canvas Osterwalder & Pigneur (2011)	◕	◑	◕	◑	◑	◑	◔	◑	◑	◑	○	○	◕	◔	◔	◔	◔
6	The new business model innovation process (BMIP) Adrodegari et al. (2017)	◕	◕	◑	◔	◑	◕	◔	◑	◑	◑	○	○	◔	○	◑	◑	◔
7	Monetarisierungs-Canvas (MTC) Fischer et al. (2023)	◔	◔	◔	◔	◔	◔	◕	◑	◔	◔	◔	◕	●	◔	◔	◔	◔
8	Data-Driven Innovation Framework Zillner (2021)	●	●	◕	◑	◑	◕	◔	◕	◕	○	◔	◔	◑	●	◔	◔	◑
9	Business Model Navigator Gassmann et al. (2021)	◑	◑	◔	○	◑	◑	◔	◔	◑	○	○	○	◔	○	○	○	○
10	Equipment-as-a-Service Canvas Kett et al. (2023)	◕	◕	●	●	◕	●	◕	◕	●	◕	◕	◕	◕	◕	◑	◑	●
	Fullfield:	○ None	◔ partial	◑ moderate	◕ advanced	● complete												

Figure 4: Framework comparison

Appendix 3: ProcessFlow TriProGenix

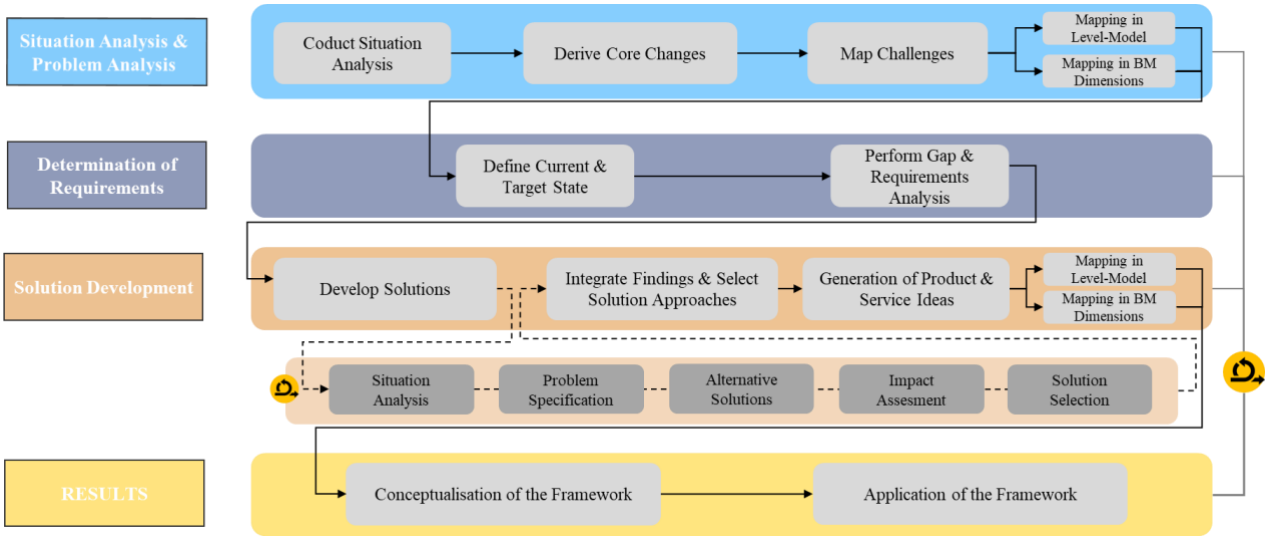


Figure 5: Processflow TriProGenix

Appendix 4: Detailed ProGenLab X

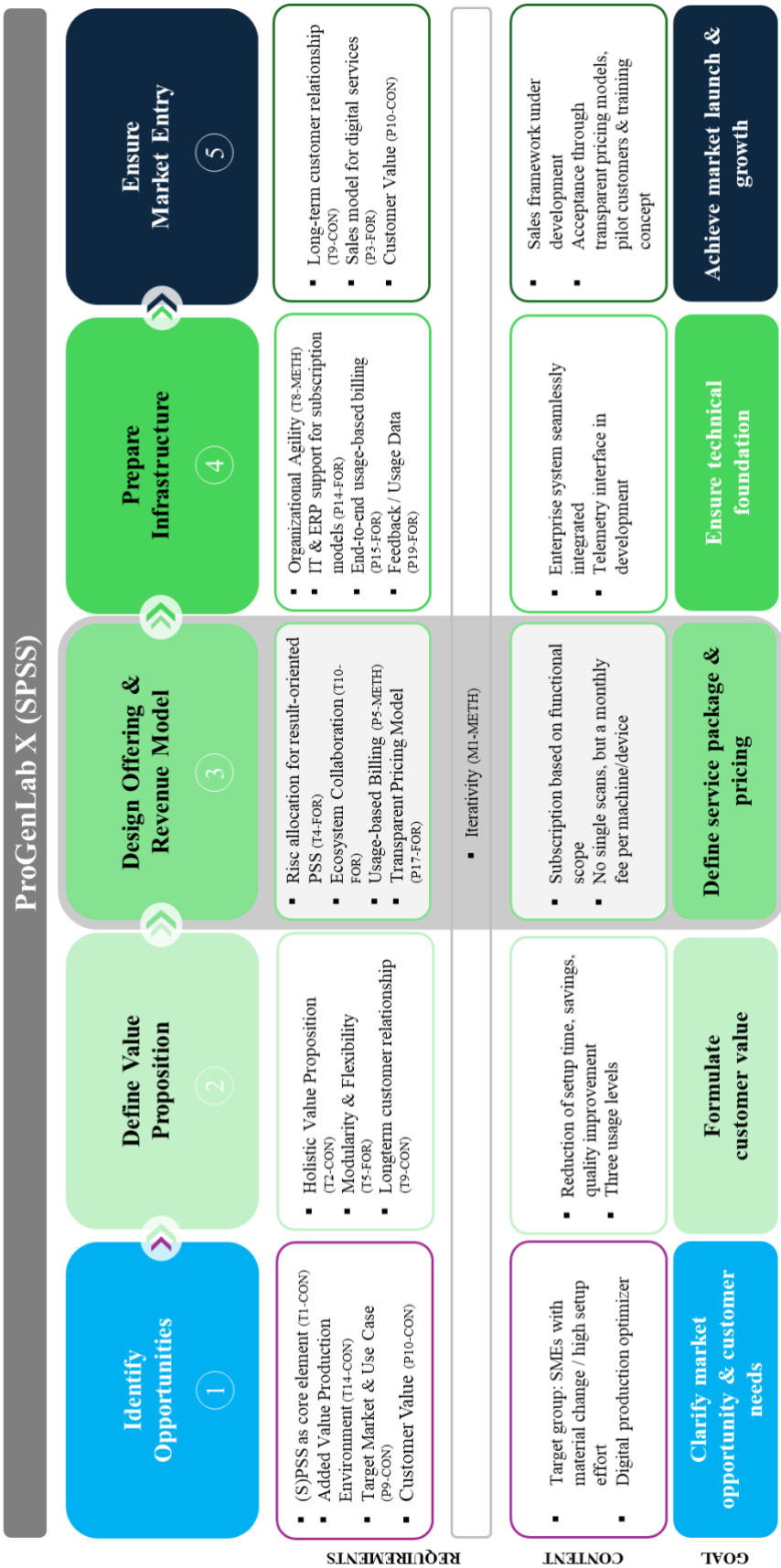


Figure 6: ProGenLab X

References

[1] Scholle, K., 2024. Beschäftigte im Maschinebau in Deutschland. https://de.statista.com/statistik/daten/studie/30821/umfrage/beschaeftigte-im-maschinenbau-in-deutschland/#:~:text=Im%20Jahr%202023%20waren%20rund%20955.000%20. Accessed 3 June 2025.

[2] Scholle, K., 2024. Umsatz im deutschen Maschinenbau. https://de.statista.com/statistik/daten/studie/3777/umfrage/umsatz-im-deutschen-maschinenbau-seit-1991/. Accessed 3 June 2025.

[3] Chen, Y., Wu, Z., Yi, W., Wang, B., Yao, J., Pei, Z., Chen, J., 2022. Bibliometric Method for Manufacturing Servitization: A Review and Future Research Directions. Sustainability 14, 1–26.

[4] Humbeck, P., Vock, E., Bauernhansl, T., 2019. Towards the Management of the Development of Product-Service Systems in Business Ecosystems - State-of-the-Art, in: 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Macao, China. 15.12.2019 - 18.12.2019. IEEE, pp. 566–570.

[5] Kett, H.J., Evcenko, D., Falkner, J., 2022. Equipment-as-a-Service (EaaS): Canvas zur Entwicklung von EaaS-Geschäftsmodellen. Fraunhofer IAO, 45 pp.

[6] Kowalkowski, C., Gebauer, H., Oliva, R., 2017. Service growth in product firms: Past, present, and future. Industrial Marketing Management 60, 82–88.

[7] Lamperti, S., Cavallo, A., Sassanelli, C., 2023. Digital Servitization and Business Model Innovation in SMEs: A Model to Escape From Market Disruption. IEEE Trans. Eng. Manage., 1–15.

[8] Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., Baines, T., 2019. Digital servitization business models in ecosystems: A theory of the firm. Journal of Business Research 104, 380–392.

[9] Lee, J., Kao, H.-A., Yang, S., 2014. Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. Procedia CIRP 16, 3–8.

[10] Sjödin, D., Parida, V., Kohtamäki, M., Wincent, J., 2020. An agile co-creation process for digital servitization: A micro-service innovation approach. Journal of Business Research 112, 478–491.

[11] Zillner, S., 2021. Business Models and Ecosystem for Big Data, in: Curry, E., Metzger, A., Zillner, S. (Eds.), The Elements of Big Data Value. Foundations of the Research and Innovation Ecosystem. Springer International Publishing, Cham, Switzerland, pp. 269–288.

[12] Ferreira, F.N.H., Proença, J.F., Spencer, R., Cova, B., 2013. The transition from products to solutions: External business model fit and dynamics. Industrial Marketing Management 42 (7), 1093–1101.

[13] Ferreira, J.J., Fernandes, C.I., Ferreira, F.A., 2019. To be or not to be digital, that is the question: Firm innovation and performance. Journal of Business Research 101, 583–590.

[14] Shen, L., Sun, C., Ali, M., 2021. Role of Servitization, Digitalization, and Innovation Performance in Manufacturing Enterprises. Sustainability 13 (9878), 1–17.

[15] Favoretto, C., Mendes, G.H.S., Oliveira, M.G., Cauchick-Miguel, P.A., Coreynen, W., 2022. From servitization to digital servitization: How digitalization transforms companies' transition towards services. Industrial Marketing Management, 104–121.

[16] Gassmann, O., Frankenberger, K., Choudury, M. (Eds.), 2021. Geschäftsmodelle entwickeln: 55+ innovative Konzepte mit dem St. Galler Business Model Navigator, 3., überarbeitete und erweiterte Auflage ed. Hanser, München.

[17] Oliva, R., Kallenberg, R., 2003. Managing the transition from products to services. International Journal of Service Industry Management 14, 160–172.

[18] Vandermerwe, S., Rada, J., 1988. Servitization of business: Adding value by adding services. European Management Journal 6 (4), 314–324.

[19] Wirth, J., Erb, S.V., Hoffmann, F., Metternich, J., Bauernhansl, T., 2024. Framework for the Introduction of Data-driven and Service-oriented Business Models in Mechanical and Plant Engineering Companies. 57th CIRP Conference on Manufacturing Systems 2024 (CMS 2024) 2024, 1–6.

[20] Kowalkowski, C., Gebauer, H., Kamp, B., Parry, G., 2017. Servitization and deservitization: Overview, concepts, and definitions. Industrial Marketing Management 60, 4–10.

[21] Verma, R., Fitzsimmons, J., Heineke, J., Davis, M., 2002. New issues and opportunities in service design research. J of Ops Management 20 (2), 117–120.

[22] Matusek, M., 2023. Exploitation, Exploration, or Ambidextrousness--An Analysis of the Necessary Conditions for the Success of Digital Servitisation. Sustainability 15 (1), 1–20.

[23] Ostrom, A.L., Bitner, M.J., Brown, S.W., Burkhard, K.A., Goul, M., Smith-Daniels, V., Demirkan, H., Rabinovich, E., 2010. Moving Forward and Making a Difference: Research Priorities for the Science of Service. Journal of Service Research 13 (1), 4–36.

[24] Teece, D.J., 2010. Business Models, Business Strategy and Innovation. Long Range Planning 43 (2), 172–194.

[25] Marcon, É., Le Dain, M.-A., Frank, A.G., 2022. Designing business models for Industry 4.0 technologies provision: Changes in business dimensions through digital transformation. Technological Forecasting and Social Change 185, 1–15.

[26] Westerlund, M., Leminen, S., Rajahonka, M., 2014. Designing Business Models for the Internet of Things 4 (7), 5–14.

[27] Chen, Y., Visnjic, I., Parida, V., Zhang, Z., 2021. On the road to digital servitization - The (dis)continuous interplay between business model and digital technology. International Journal of Operations & Production Management 41 (5), 694–722.

[28] Frank, A.G., Mendes, G.H., Ayala, N.F., Ghezzi, A., 2019. Servitization and Industry 4.0 convergence in the digital transformation of product firms: A business model innovation perspective. Technological Forecasting and Social Change (141), 341–351.

[29] Paiola, M., Gebauer, H., 2020. Internet of things technologies, digital servitization and business model innovation in BtoB manufacturing firms. Industrial Marketing Management 89, 245–264.

[30] Hein, A., Schreieck, M., Riasanow, T., Setzke, D.S., Wiesche, M., Böhm, M., Krcmar, H., 2019. Digital platform ecosystems 1, 12 pp.

[31] Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a theory of ecosystems. Strategic Management Journal 39 (8), 2255–2276.

[32] Koch, M., Polst, S., Naab, M., Bartels, N.A., Storck, S., Trapp, M., Rauch, B., Jeswein, T., 2023. Digitale Ökosysteme in Deutschland.

[33] Pidun, U., Reeves, M., Schüssler, M., 2019. Do you need a Business Ecosystem? Accessed 12 December 2023.

[34] Gawer, A., Cusumano, M.A., 2013. Industry Platforms and Ecosystem Innovation. Journal of Product Innovation Management 31 (3), 417–433.

[35] Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a Theory of Ecosystems. SSRN Journal.

[36] Kolagar, M., Parida, V., Sjödin, D., 2022. Ecosystem transformation for digital servitization: A systematic review, integrative framework, and future research agenda. Journal of Business Research 146, 176–200.

[37] Neely, A., 2008. Exploring the financial consequences of the servitization of manufacturing. Oper Manag Res 1 (2), 103–118.

[38] Porter, M.E., Heppelmann, J.E., 2014. How smart, connected products are transforming competition. Harvard business review 92 (11).

[39] Tukker, A., 2004. Eight types of product–service system: eight ways to sustainability? Experiences from SusProNet. Bus Strat Env 13 (4), 246–260.

[40] Huikkola, T., Kohtamäki, M., Ylimäki, J., 2022. Becoming a smart solution provider: Reconfiguring a product manufacturer's strategic capabilities and processes to facilitate business model innovation. Technovation 118, 1–13.

[41] Mathieu, V., 2001. Service strategies within the manufacturing sector: benefits, costs and partnerships. International Journal of Service Industry Management 12 (5), 451–475.

[42] Schneider, T., Braig, B., Junker, J.-A., Keese, C., 2023. Umdenken as a service: Die Zukunft des Maschinenbaus in Europa : wie wir Ressourcen effizient nutzen und Wertschöpfung steigern, 1. Auflage ed. hy The Axel Springer Consulting Group, Berlin, 125 pp.

[43] Schüritz, R., Satzger, G., 2016. Patterns of Data-Infused Business Model Innovation, in: 2016 IEEE 18th Conference on Business Informatics (CBI). 2016 IEEE 18th Conference on Business Informatics (CBI), Paris, France. 29.08.2016 - 01.09.2016. IEEE, pp. 133–142.

[44] Hevner, A., 2007. A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19 (2), 87–92.

[45] Hevner, A., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research. MIS Quarterly 28 (1), 75–105.

[46] vom Brocke, J., Hevner, A., Maedche, A., 2009. Introduction to Design Science Research, 1–13.

[47] Misoch, S., 2019. Qualitative Interviews: 2., erweiterte und aktualisierte Auflage. Walter de Gruyter GmbH, Berlin, Boston, 320 pp.

[48] Wirth, J., Hermann, M., Bauernhansl, T., 2025. Navigating towards New Businesses: An Implementation Framework for Data-Driven and Service-Oriented Business Models. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[49] Wirth, J., Kallen, P., Bauernhansl, T., 2025. Adaptation and Operationalization of a Process Model for the Introduction of Service-Oriented and Data-Driven Business Models in Mechanical and Plant Engineering. Procedia CIRP 134, 384–389.

[50] Wirth, J., Seidel, T., Bauernhansl, T., 2025. Development Of A Framework For A Structural Analyzation And Examination Regarding The Need Of Serviceoriented And Datadriven Business Models In Mechanical- And Plant-Engineering. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[51] Albers, A., Burkhardt, N., Meboldt, M., Saak, M., 2005. SPALTEN Problem Solving Methodology in the Product Development, in: Engineering Design and the Global Economy: 15th International Conference on Engineering Design - ICED 2005, Melbourne, 15. - 18. August 2005. Hrsg.: A. Samuel. The Design Society, p. 90671.

[52] Gregor, S., Hevner, A.R., 2013. Positioning and Presenting Design Science Research for Maximum Impact. MISQ 37 (2), 337–355.

[53] Pohl, K., Rupp, C., 2021. Basiswissen Requirements Engineering: Aus- und Weiterbildung nach IREB-Standard zum Certified Professional for Requirements Engineering. dpunkt.verlag, [S.l.], 1 online resource.

[54] Li, H., Zhao, J., Cao, Y., Su, L., Zhao, Z., Zhang, Y., 2024. Servitization and product service system: A literature review on value creation. Technological Forecasting and Social Change 208, 123724.

[55] Aas, T.H., Breunig, K.J., Hellström, M.M., Hydle, K.M., 2020. Service-oriented business models in manufacturing in the digital era: Toward a new taxonomy. International Journal of Innovation Management 24 (08), 1–15.

[56] Sjödin, D., Parida, V., Visnjic, I., 2022. How Can Large Manufacturers Digitalize Their Business Models? A Framework for Orchestrating Industrial Ecosystems. California Management Review 64 (3), 49–77.

[57] Silva, H., Soares, A.L., 2021. Digital Platforms as Enablers of Smart Product-Service Systems, in: Camarinha-Matos, L.M., Boucher, X., Afsarmanesh, H. (Eds.), Smart and Sustainable Collaborative Networks 4.0, vol. 629. Springer International Publishing, Cham, pp. 506–513.

[58] Osterwalder, A., Pigneur, Y., 2011. Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus Verlag.

[59] Kohtamäki, M., Parida, V., Patel, P.C., Gebauer, H., 2020. The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. Technological Forecasting and Social Change 151.

[60] Aurich, J.C., Fuchs, C., Wagenknecht, C., 2006. Life cycle oriented design of technical Product-Service Systems. Journal of Cleaner Production 14 (17), 1480–1494.

[61] Müller, P., Kebir, N., Stark, R., Blessing, L., 2009. PSS Layer Method – Application to Microenergy Systems, in: Sakao, T., Lindahl, M. (Eds.), Introduction to Product/Service-System Design. Springer London, London.

[62] Adrodegari, F., Pashou, T., Saccani, N., 2017. Business Model Innovation: Process and Tools for Service Transformation of Industrial Firms. Procedia CIRP 64, 103–108.

[63] Fischer, S., Golz, P., Greven, D., Herr, C., 2023. Monetarisierung plattformbasierter Produkte und Services im Maschinen- und Anlagenbau.

Biography

Jonas Wirth (*1993) has been a doctoral candidate at Fraunhofer IPA/IFF, of the University of Stuttgart, in cooperation with TRUMPF SE + Co. KG since 2022. His research interest is the scale-up of service-oriented business models in value networks. In the research project Factory-X, Mr. Wirth is responsible for the business model project.

Dennis Unic (*1991) is a master student of Industrial Management at the University of Aalen and has conducted research in collaboration with TRUMPF SE + Co. KG about service-oriented and data-driven business models.

Thomas Bauernhansl (*1969) has been Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since September 2011. He is also Director of the Institute for Industrial Manufacturing and Management IFF at the University of Stuttgart. His scientific focus is on digital and biological transformation, production organization, factory planning, automation, and robotics.

8th Conference on Production Systems and Logistics

From Uncertainty to Value: The Scoreboard and Feasibility Gate for Early-Stage Financial Planning in Industrial Data Ecosystems

Jonas Wirth[1,3,4], Fabian Meldt[2], Felix Hoffmann[2], Matthias Weigold[2], Thomas Bauernhansl[1,4]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12, 70569 Stuttgart, Germany*
[2] *PTW of TU Darmstadt*
[3] *TRUMPF, Johann-Maus-Straße 2, 71254 Ditzingen, Germany*
[4] *Institute of Industrial Manufacturing and Management (IFF), University of Stuttgart*

Abstract

Manufacturing is shifting from products to service-led offerings. Extending this trajectory, firms create value across organizational boundaries via industrial data ecosystems (e.g. Factory-X). These ecosystems promise efficiency, revenue, and sustainability, yet managers lack firm-level guidance on the economic feasibility of participation under federated collaboration. Due to the lack of decision-support artefacts for assessing a company's financial benefit from participating in industrial data ecosystems, this paper develops an early-stage financial planning artefact. It reuses business-model decomposition from prior evaluation approaches but extends it with multi-role scoring and TCE-based feasibility reasoning for federated industrial data ecosystems. Following a Design Science Research approach, including expert input, the artefact consists of a qualitative Scoreboard and a TCE-based Feasibility Gate, complemented by a prototype semi-quantitative bridge. The Scoreboard captures role-agnostic drivers across value proposition, value creation, value capture, ecosystem network, and governance, including collaboration frictions and transaction-cost factors. A prototype semi-quantitative bridge maps weighted driver scores to scenario parameters and cash-flow templates, yielding indicative metrics such as ROI, NPV, and payback. The artifact is demonstrated in the Factory-X context, where one use case and expert workshops provide evidence of usability and decision relevance. The model supports go/no-go and prioritization decisions and clarifies added value and data-sharing requirements. Within the broader TriProGenix research framework aimed at enabling service-oriented and data-driven business models in mechanical and plant engineering, the proposed model is positioned at the ecosystem layer and complements established approaches for business model innovation at the product and service levels. For research, the work shows how qualitative business-model drivers can be structured as early-stage feasibility signals for subsequent financial planning in federated ecosystems, while respecting transaction costs (TC). For practice, it offers a tool that helps companies assess ecosystem feasibility, weigh role-specific trade-offs, and attribute added value to ecosystem participation.

Keywords

Economics, Decision making, Industrial data ecosystem, Data-driven, Business Model Innovation, TriProGenix

DOI: https://doi.org/10.15488/20997
ISSN: 2701-6277

1. Introduction/Motivation

Machinery and plant engineering faces persistent competitive pressure to transition from one-time equipment sales toward data-driven, service-oriented value creation [1,2]. Digital Twins and Industrial Internet of Things enable new service models, ranging from predictive maintenance to performance-based contracts, which are promising recurring revenue and enhanced competitiveness [3]. Yet despite technical maturity, data driven service models rarely scale beyond pilots [4]. The barrier seems not to be technological but economic [5]. Companies often lack structured frameworks to evaluate whether data ecosystem participation generates positive returns. This assessment gap becomes critical under fundamental uncertainty where historical data is absent. Traditional evaluation methods prove inadequate for forecasting in early-stage ecosystems [6]. Empirically, this uncertainty translates into persistent underutilization [7]. The European Commission estimates around 80% of collected industrial data in Europe to remain unused, not because of technical limitations but because of unresolved economic questions. [8]. Data sovereignty concerns, undefined monetization models, and uncertain governance overhead cause firms to treat data sharing as risk rather than opportunity [4,9,10]. External pressures simultaneously intensify the urgency of addressing this gap. The EU Data Act now defines a mandatory framework for data access rights and obligations that firms cannot ignore, thereby reshaping commercial incentives in data ecosystems [11]. Simultaneously, global competitors in the United States and China increasingly monetize industrial data through vertically integrated platforms, creating competitive disadvantage for European firms without economically viable data ecosystems [12,13]. Manufacturing-X and Factory-X respond by establishing sovereign technical and governance foundations, yet they leave unanswered the critical question [14,15]: Under what conditions should a firm actually participate? This paper addresses this research gap by introducing an early-stage financial planning artefact that translates business-model drivers and governance logic into structured factors for assessing the economic viability of participation in federated industrial data ecosystems. The artefact consists of a qualitative Scoreboard and a TCE-based Feasibility Gate. It operationalises business-model components for decentralised architectures [16] and explicitly incorporates transaction-cost effects. Rather than forecasting precise financial outcomes, it provides decision-makers with a rigorous evaluative logic for assessing data-driven opportunities under uncertainty within the data-ecosystem context. Situated within the authors' broader TriProGenix research programme on implementing service-oriented and data-driven business models in digital ecosystems, the artefact contributes an ecosystem-level feasibility module, not by introducing fundamentally new constructs, but by integrating and operationalising existing business-model and TCE perspectives for role selection and participation strategies in Factory-X-type environments.

2. Fundamentals

2.1 Business Models and Servitization

Business models provide a structured representation of how an organisation creates, delivers and captures value [17–19]. For analysis, this logic can be decomposed into four components: value proposition, value creation, value network and value capture [16]. Canonical frameworks such as the Business Model Canvas and the St. Gallen Business Model Navigator operationalise these components and highlight the interdependence between "Who", "What", "How" and "Value" in business-model design [20,21].

Servitization captures the shift from product-centric to service-oriented value creation [22,23]. In machinery and plant engineering, commoditised hardware, regulation and outcome-oriented customer expectations make this shift increasingly mandatory [24]. Product-Service Systems (PSS) integrate products and services into lifecycle-oriented solutions [25–28], with result-oriented PSS tying payment to outcomes and redistributing risk between provider and customer [29,27]. Smart PSS extend this logic by embedding IoT connectivity, cloud platforms and analytics into offerings, making data the central resource for monitoring, optimisation and AI-based assistance [30,31]. This evolution creates the basis for data-driven, usage- and performance-based business models in industrial production [14,32,7].

2.2 Data Ecosystems

Industrial value creation increasingly depends on secure multilateral data exchange rather than isolated analytics. Data ecosystems are socio-technical systems in which autonomous actors exchange data via standardised interfaces under shared policies and trust frameworks [33]. Federated data ecosystems allow participants to retain sovereignty over their data while coordinating through common standards [34]. Data ecosystems typically involve data providers, data users, and enabling actors that provide infrastructure and trust services [35,36]. From a transaction-cost economics perspective, economic viability requires governance that keeps ex-ante and ex-post coordination, negotiation and compliance costs manageable relative to captured value [37]. Gaia-X and Factory-X exemplify this federated logic by combining sovereign data sharing, interoperability and trust mechanisms with decentralised infrastructure for machinery and plant engineering [15,38,39]. In such settings, business-model innovation is typically focused and data-centric, centred on new ways to monetise and share operational data across organisational boundaries [40,41].

3. Related Research

3.1 Model Requirements

This paper consolidates theoretical and practical requirements for a decision-grade evaluation model for participation in federated industrial data ecosystems and uses them as a reference point to assess existing approaches. The requirements fall into three groups. Content-related requirements, derived from business-model, servitization, data-ecosystem and TCE research, include value proposition, value creation, value capture, value network, ecosystem roles, and transaction-cost and governance aspects [22,42,25,19]. Methodological requirements emphasise role specificity, scenario-based reasoning, and decision support that can combine qualitative and quantitative evidence, for example through anchored scales and weighting schemes [43,44]. Formal requirements comprise transparency, auditability and workshop fit, together with a clear bridge to cash-flow-based financial indicators for later project stages. Insights from Factory-X and TRUMPF confirm these requirements and add early-stage applicability, manageable effort and comparability across heterogeneous roles as central design criteria for an evaluation model.

3.2 Related Approaches

Related work can be grouped into three streams. The first covers business-model and servitization frameworks and process models. Generic approaches such as the Business Model Canvas, the Magic Triangle and BM-Dimensions structure value propositions, revenue logic, key resources and partner networks and are widely used to design and communicate business models [45,20,21]. Domain-specific contributions on smart product-service systems and Equipment-as-a-Service extend this logic to lifecycle-oriented, data-enabled service offerings in manufacturing [46–48]. Recent work in mechanical and plant engineering, in particular the TriProGenix framework, provides a structured transformation logic for data-driven and service-oriented business models [49]. These approaches offer strong conceptual guidance on value propositions, service design and organisational change. However, they do not provide an early-stage financial assessment of ecosystem participation or a TC-aware role evaluation for federated industrial data spaces. The present work therefore extends the TriProGenix logic with a dedicated feasibility module for Factory-X-type ecosystems.

The second stream covers ecosystem and TCE-based perspectives. Platform and industrial data ecosystem research discusses roles, governance and value co-creation in federated data spaces and highlights the role of ex ante and ex post transaction costs for economic feasibility [50,34,51]. Work in the Gaia-X and Manufacturing-X context stresses sovereign data sharing, interoperability and trust mechanisms as foundations for cross-firm service models [38,39]. Formal analyses of ecosystem feasibility use TCE to express viability as a condition where additional value must at least compensate additional transaction costs and where governance structures mitigate hazards and coordination effort [37]. These contributions sharpen

the understanding of governance design and feasibility at ecosystem and use-case level. However, they remain largely conceptual and do not offer a practical role-level instrument that links transaction-cost drivers with concrete business-model options.

The third stream comprises monetary evaluation approaches for data-driven offerings. Classic investment appraisal, ROI and NPV-based methods and recent "value of data" contributions provide tools to quantify expected cash flows, risks and returns, but are usually agnostic to ecosystem roles and federated data spaces [52–54]. The method proposed by Hoffmann is particularly relevant as a reference point. It develops a practice-oriented evaluation approach for data-based business models by decomposing them into value proposition, value capture, value creation and value network and aggregating the results through investment calculation [55]. This logic presupposes a single-provider perspective, predominantly bilateral customer relationships and a project context in which key revenue and cost parameters can be estimated with reasonable confidence. In early-stage Factory-X-type ecosystems, roles, transaction patterns and pricing mechanisms are still fluid, while transaction and governance costs are difficult to allocate to individual bilateral relationships. A direct transfer of Hoffmann's quantitative investment logic would therefore risk spurious precision [55]. This paper reuses Hoffmann's business-model decomposition as a conceptual backbone but differs in three respects. It targets federated multi-role settings rather than a single-provider context, treats governance and transaction costs as explicit feasibility drivers, and uses qualitative scoring as an early-stage decision device rather than direct investment calculation. The gap is therefore not the absence of business-model decomposition, but the absence of an early-stage, role-based feasibility instrument for federated ecosystems under governance and transaction-cost uncertainty.

Appendix B maps the three streams of related work to the requirement profile from Section 3.1. It shows that no existing contribution fully meets the combined requirements from theory and practice. Business-model and servitization frameworks structure content and transformation paths but lack transaction-cost-aware feasibility checks for specific roles in federated industrial data ecosystems [56]. Ecosystem- and TCE-based work sharpens understanding of governance choices and feasibility conditions yet remains largely conceptual and does not provide participant-level decision support [37]. Monetary evaluation approaches, including Hoffmann's, offer financial depth and a link between business-model design and investment calculation, but are tied to single-firm settings and do not operationalise multi-role, federated data spaces with uncertain transaction and governance costs [55]. These gaps motivate the Scoreboard and Feasibility Gate as an early-stage evaluation artefact that systematically combines business-model drivers and TCE factors for federated, role-based participation decisions.

4. Methodology

4.1 Design Science Research

This paper adopts a Design Science Research (DSR) approach to develop and evaluate the Scoreboard and Feasibility Gate for early-stage financial planning in decentralized industrial data ecosystems, following Hevner's three-cycle view and the Design Science Research Methodology (DSRM) [57–59]. DSR links artifact design to both theory and practice. The process follows the knowledge, environment, and design cycles and covers problem identification, objective definition, artifact design, demonstration, and evaluation [59]. The knowledge base is established through a structured literature review following vom Brocke et al. and a concept-centric synthesis in the sense of Webster and Watson [60,61]. A search string derived from the research questions covers business models and servitization, data ecosystems, evaluation methods and machinery and plant engineering. It is applied to four databases with aligned Boolean logic. After duplicate removal and title and abstract screening, the remaining publications are analysed in depth and organised in a concept matrix covering the four domains. Backward snowballing extends the corpus and, together with

practice-oriented sources, provides the theoretical and content requirements for the Scoreboard [60–62]. Figure 1 summarises the search and selection process.

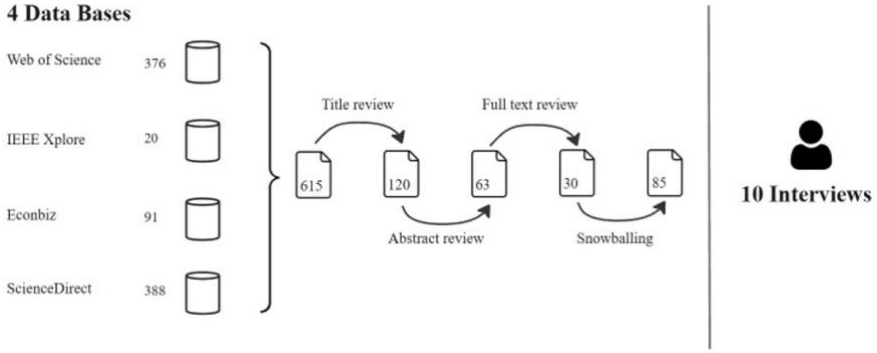


Figure 1: Structured literature research, according to VOM BROCKE et al. [60] and Interview validation.

The environment and relevance cycle are instantiated through the Factory-X initiative and its industrial partners, in particular TRUMPF, with Use Case 2.4 "Condition-Monitoring-Led Services" as the application scenario analysed in Section 6 [57–59]. Experts from component suppliers, machine builders, factory operators and software or infrastructure providers inform the design cycle. The resulting TriProGenix-based framework [49] and the Scoreboard, including anchored 1–5 scales, evidence and confidence tags, optional AHP-based weighting and the TCE-based Feasibility Gate, are detailed in Sections 4.2 and 5, while their application and validation are reported in Section 6. In simple terms, the evaluation logic proceeds in three steps: experts score the five dimensions, the scores are aggregated into indices, and the Feasibility Gate interprets the resulting benefit-cost pattern.

4.2 TriProGenix

Building on this DSR foundation, the paper uses the TriProGenix (TPG) framework from Wirth as a reference model for transformation towards data-driven, service-oriented and ecosystem-based business models in mechanical and plant engineering [63,49]. TPG combines a navigation system for structured change with the idea that firms can generate profit from products, services and ecosystem participation. It provides the conceptual umbrella under which this paper positions the Factory-X evaluation model. Within this framework, the Scoreboard and Feasibility Gate operationalise the "profit from ecosystems" dimension by making participation choices and role configurations assessable. Figure 2 gives a condensed overview of the framework and its three core dimensions: development levels, problem-solving process and method toolkit [49]. The first dimension comprises vertical development levels from normative orientation to strategic positioning and organisational implementation [49,64]. In this paper, the Scoreboard is mainly situated at the interface between strategic and organisational levels, where concrete ecosystem options must be prioritised and justified. The second dimension is a problem-solving process that structures transformation from situation analysis and requirements definition through solution development and business-model choice to implementation and scaling, drawing on generic procedures such as the SPALTEN system [63].

In our context, the Scoreboard supports requirements definition and business-model selection by making financial viability and transaction-cost effects of Factory-X roles explicit. The third dimension is a method portfolio that groups analytical and design instruments for strategy work, diagnosis, organisational analysis, decision-making and project management [49]. The proposed evaluation model extends this toolkit with a role-agnostic, transaction-cost-aware feasibility instrument for federated data ecosystems. In combination, TPG and the Scoreboard form a coherent approach that links product- and service-based profitability with ecosystem-level financial planning in Factory-X. More broadly, they provide foundational building blocks for a decision-oriented rule set and position the artefact as an early pre-investment screening step before detailed business-case calculation within the TriProGenix programme.

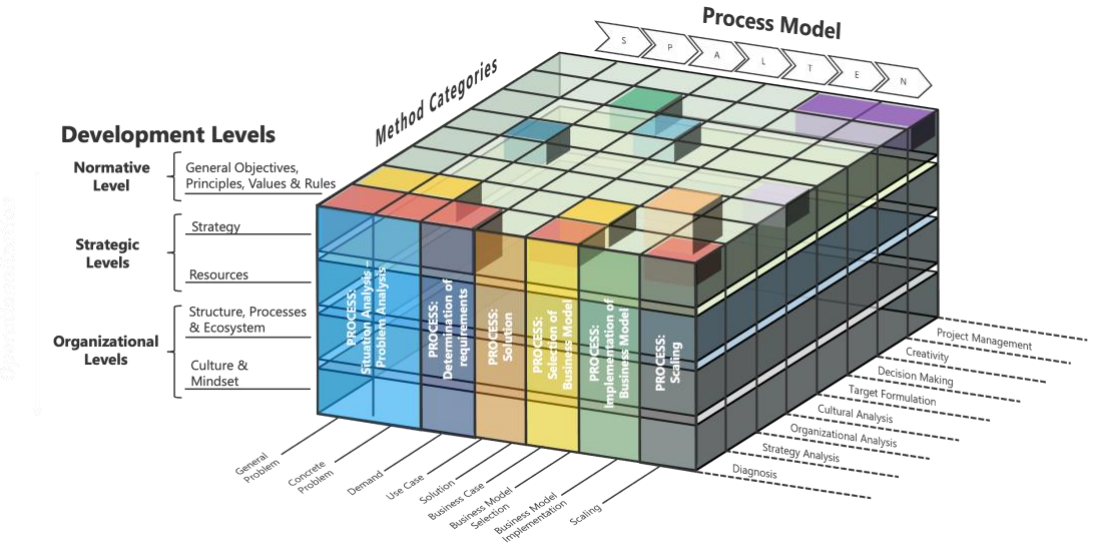


Figure 2: Overview of the TriProGenix Framework [49].

5. Data Ecosystem Framework

5.1 Data Ecosystem Pyramid Model

The model developed in this paper constitutes the primary design-science research (DSR) artefact. It translates the requirement profile outlined in Section 3 and the gaps identified in related research into a compact evaluation framework for Factory-X-type data ecosystems. Within the TriProGenix logic of profit from products, services, and ecosystems, the artefact focuses on the ecosystem layer and addresses the question of under which conditions participation is financially viable for a given role in a specific scenario, rather than prescribing how firms should design new business models [49,65]. For readability, the framework can be read in two parts: the five-dimensional Scoreboard and the Feasibility Gate.

The framework is conceptualised as a pyramid (see Figure 3). A Governance/TCE layer forms the rectangular base, capturing the transaction costs and governance mechanisms that determine whether value from Factory-X participation can be realised in practice. Resting on this base are three lateral faces representing the business-model dimensions What, How and Value, which summarise the expected value proposition, the role's capabilities to deliver it and the mechanisms for value capture and risk allocation. The apex, Ecosystem & Network, reflects the value-network perspective, encompassing the concrete role constellation, partner structure and the focal actor's position within the ecosystem. Collectively, these five dimensions integrate business-model logic with transaction-cost economics to enable a structured assessment of the attractiveness of ecosystem participation for a given role. A workshop-oriented Scoreboard operationalises the dimensions for expert workshops, while a transaction-cost-based Feasibility Gate assesses whether the aggregated benefits justify the associated transaction costs for the role.

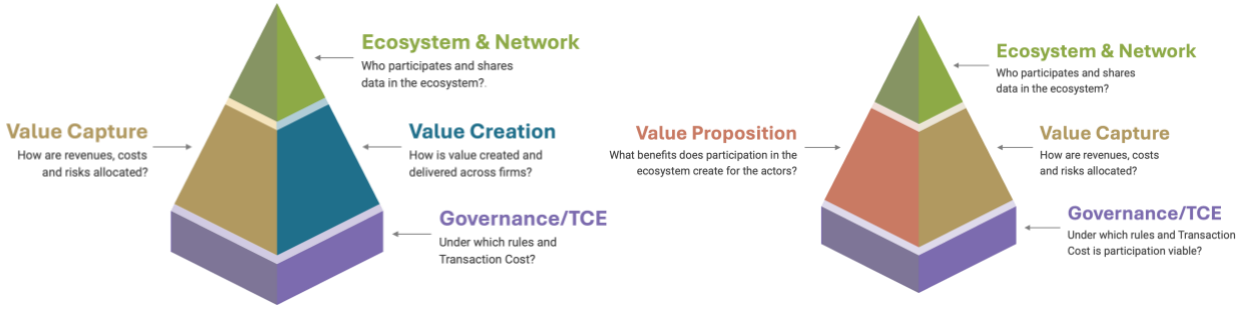


Figure 3: Data Ecosystem Pyramid model from two perspectives.

The What dimension captures whether a role can expect relevant value from participation, such as efficiency and quality gains, new services or regulatory and sustainability benefits. How reflects the ability to realise this value, including data, analytics and service capabilities and the integration of condition-monitoring and related services into existing processes. Value covers value-capture and risk-sharing aspects, including workable pricing logics for data-enabled services, alignment with internal steering and margins and the allocation of investment and operating risk. Who corresponds to the value-network building block and considers ecosystem configuration, availability of relevant partners and data sources, power and dependency structures and the actor's influence over key design choices. The Governance/TCE base anchors these views in ex-ante and ex-post transaction costs in federated data spaces, such as partner search and onboarding in a Gaia-X-style environment on the one hand and ongoing policy maintenance and dispute handling during operation on the other. These governance drivers do not create customer value on their own but determine whether additional value from the four business-model dimensions can be realised at acceptable cost.

5.2 Interview Scoreboard

In practical use, the pyramid is operationalised through a Scoreboard (Figure 4) that decomposes each dimension into a set of drivers. The practical logic is simple: first score the five dimensions, then interpret them through the Feasibility Gate. For each dimension, a fixed set of drivers translates the abstract constructs into role-independent questions for expert workshops. A filled interview example is provided in Appendix D. Drivers are scored on anchored 1-5 scales with qualitative descriptors. Evidence tags indicate whether a score rests on internal data, external sources or expert judgement, and confidence flags mark assumptions that should be treated cautiously in interpretation and sensitivity analysis. Optional weighting using additive value functions and AHP-derived weights reflect the relative importance of drivers and dimensions. Detailed driver descriptions and rubrics lie outside the scope of this paper, which concentrates on the dimension logic, aggregation and interpretation.

In the baseline configuration, anchored 1-5 driver scores are normalised to dimension indices, aggregated by simple additive value functions, and optionally reweighted through AHP for sensitivity analysis. The indices are interpreted jointly rather than in isolation, and workshop discussions explicitly address dependencies between dimensions, such as value-proposition strength increasing both benefit and coordination effort, or orchestrator-dominated network structures accelerating adoption while shifting bargaining power and risk. For example, a role with high What and Who scores but weak Value and Governance/TCE scores may appear strategically attractive yet still fail the Feasibility Gate because monetisation logic and coordination costs remain unresolved.

The Feasibility Gate condenses transaction-cost-based reasoning into a simple decision lens. It combines the business-model indices into a benefit profile and contrasts this with the Governance/TCE index. Participation in Factory-X is considered economically attractive for a role if the additional value at least compensates the additional transaction and governance costs in the considered scenario. Instead of computing net present values or detailed cash-flow projections, the Gate maps the qualitative assessment to a small set of feasibility states, such as clearly feasible, feasible under conditions or currently not feasible, together with an explanation of which dimensions act as bottlenecks or enablers. A positive or conditional result then serves as the entry point for subsequent quantitative planning, including scenario assumptions on adoption, pricing, governance-cost allocation, and cash-flow modelling. In managerial use, the artefact is intended for early pre-investment screening after role and use-case definition but before detailed budgeting or investment approval. The resulting pyramid framework connects business-model logic, data-ecosystem thinking and TCE in a single artifact that is rich enough to capture the main qualitative drivers of ecosystem value and governance effort, yet simple enough for workshop use with heterogeneous stakeholders.

Dimension	Code	Driver
Ecosystem & Network (Who?)	N-1	Role Coverage & Partner Identification
	N-2	Network-effect Potential
Value Proposition (What?)	W-1	Efficiency Gains
	W-2	Sustainability & Compliance Gains
	W-3	Insight-driven Revenue Enablement
Value Creation (How?)	H-1	Data-pipeline CAPEX/OPEX
	H-2	Analytics & Lifecycle Effort
	H-3	Inter-disciplinary Coordination Complexity
	H-4	Organisational Readiness & GTM Capability
	H-5	Interoperability Maturity
Value Capture (Value?)	V-1	Direct Monetisation Flexibility
	V-2	Net Ecosystem Rents
	V-3	Indirect Revenue Uplift
	V-4	Market-entry Discounts & Incentives
Governance/TCE	G-1	Ex-ante Transaction Costs
	G-2	Ex-post Transaction Costs
	G-3	Sovereignty & Compliance Overhead

Figure 4: Scoreboard structure with the 17 drivers assigned to the five dimensions.

6. Application and Validation

This section demonstrates and validates the artefact in Factory-X Use Case 2.4, Condition-Monitoring-Led Services (CMLS), and summarises the scoring setup, role coverage, and main feasibility results. CMLS connects factory operators, machine builders and component suppliers through a software-provider infrastructure [15]. The scenario revolves around remote condition monitoring of machinery using sensor data to enable preventive and predictive maintenance and decision support for maintenance scheduling. Operators expect higher availability and lower total cost of ownership, machine builders see usage-based revenue streams and closer customer relationships, and component suppliers gain fleet-level insights for root-cause analysis and reliability validation. The use case remains in pilot phase, with substantial integration effort per customer and unresolved questions around liability, intellectual-property exposure and governance overhead. This combination of clear value hypotheses and pronounced transaction-cost frictions makes CMLS a suitable test case for the Scoreboard and Feasibility Gate and guided the DSR-style demonstration [57]. From a TriProGenix perspective, CMLS also constitutes a critical test case for the ecosystem-profit dimension of the framework, as it exposes the conditions under which ecosystem narratives fail to materialise as economically viable roles for participating firms.

Ten experts from machinery and plant engineering who are active in the use case participated in interview and scoring sessions, covering the main Factory-X roles in CMLS (three machine builders, three component suppliers, three factory operators and one software provider). Each session opened with a short briefing on the use case, the pyramid dimensions and the interpretation of the anchored 1–5 scales. Experts then scored the seventeen drivers of the Scoreboard for the role under consideration, provided brief qualitative justifications and tagged whether their assessment drew primarily on internal experience or external evidence. Confidence flags down-weighted scores based on weak or speculative information. The sessions produced complete role-specific Scoreboard sheets with driver scores for each of the five dimensions, evidence annotations and comments on perceived bottlenecks and enablers in the Factory-X setting. For each role and scenario, driver scores were normalised and aggregated into indices for What, How, Value, Who and Governance/TCE using simple additive value functions [43]. These indices were then combined into a qualitative benefit profile and contrasted with the Governance/TCE index [37], yielding feasibility states per role that indicate where participation in CMLS appears clearly attractive, where it is conditionally attractive and where governance overhead or missing business-model clarity currently block participation. Sensitivity analyses tested the effect of AHP-based weighting of drivers and dimensions and alternative normalisation choices [44]. Across tested configurations, the qualitative pattern of strong and weak dimensions remained stable, suggesting that the framework does not rely on a single arbitrary parameterisation.

The workshop application provides artifact-level evidence of usefulness and transparency. The substantive feasibility results, however, remain specific to the examined Factory-X use case and should not be read as generalisable performance claims across industrial data ecosystems. Experts reported that the pyramid and the Scoreboard structure helped them distinguish between value-side and governance-side arguments and to articulate trade-offs between roles that are usually only discussed informally. The anchored scales were judged concrete enough for workshop use in the examined Factory-X setting. Several experts also emphasised that the Feasibility Gate highlights governance bottlenecks more clearly than revenue potential. They also pointed to a clear next step, namely translating positive or conditional feasibility signals into scenario assumptions for pricing, adoption, governance-cost allocation, and later financial projections once Factory-X pricing schemes and contractual templates mature further. Overall, the CMLS application shows that the Scoreboard and Feasibility Gate can be applied in a realistic Factory-X context with heterogeneous roles and incomplete information. The artifact produces auditable, role-specific assessments of ecosystem readiness and feasibility without requiring full business-case calculations and helps prioritise where further design effort in Factory-X should focus from a financial-planning perspective.

7. Results and Outlook

7.1 Scoring Results

The application in Factory-X Use Case 2.4 yields three central findings. First, the Feasibility Gate shows that the use case is not economically feasible for any of the examined roles under current conditions. For all ten participants, the combined benefit index derived from the business-model dimensions remains below the TC index from the Governance/TCE base, so participation fails the condition $B_i \geq TC_i$ for component suppliers, machine builders, factory operators and the software-provider role alike. The welfare check at ecosystem level is negative as well, because the sum of role-specific benefit–cost differences is clearly below zero [37]. This pattern is robust across equal-weight and AHP-based weighting schemes and confidence-based sensitivity tests and is characteristic of pre-critical-mass ecosystems where external interventions are required before voluntary participation becomes rational [66,67]. It therefore points to structural feasibility barriers in the current Factory-X design rather than to artifact-specific effects of the evaluation model. Figure 5 (see Appendix C) summarises the role-specific benefit and TC indices and the resulting Feasibility Gate outcomes.

Second, the Scoreboard shows how these barriers differ across roles and which qualitative drivers dominate the TC index. Factory operators perceive the strongest benefit from participation because they expect higher availability, better maintenance planning and reduced total cost of ownership, but face high governance and integration burden, especially coordination with multiple suppliers and compliance with evolving data-sovereignty requirements. Machine builders score high on network potential but struggle to articulate sustainable value-capture logics in a federated setting, which depresses their Value dimension and leaves the overall benefit index weak. Component suppliers report the lowest benefit and face the highest relative drag from sovereignty and compliance overhead; the Governance/TCE lane, and specifically the driver capturing credential management, audit requirements and multi-party compliance, emerges as the dominant bottleneck for this role. The software provider benefits from strategic positioning but faces concentrated liability and intellectual-property risk. Across all roles, governance-related drivers consistently receive very low scores, while value-proposition and network potential remain moderate to high, confirming that the main barrier lies in the transaction-cost architecture rather than in a lack of perceived opportunity.

Third, the results highlight a pronounced readiness gap between ecosystem narratives and organisational and technical capacity. On average, network-related scores markedly exceed value-creation scores, indicating that participants see clear strategic potential in Factory-X but doubt their ability to deliver and capture value in the short term. Experts point to fragmented infrastructure, immature interoperability standards and limited

in-house analytics capabilities as reasons why the How dimension lags the perceived opportunity in What and Who. This gap is most visible for small and medium-sized component suppliers and machine builders with limited IT staffing, while larger manufacturers show higher readiness indices but still fail the Feasibility Gate because of governance overhead. The Scoreboard thus makes explicit that, in this early-stage and pre-critical-mass Factory-X setting, execution risk and TC risk jointly dominate classical market or technology risk.

7.2 Implications

These findings have several implications for research and practice. Conceptually, they support the view that TCE is a useful lens for ecosystem feasibility when combined with a role-level business-model perspective in federated data spaces. The pyramid framework operationalises this link by translating qualitative business-model and governance drivers into structured indices and a feasibility condition that can be applied in workshop settings without detailed cash-flow projections. It contributes a role-agnostic, scenario-specific evaluation instrument for early-stage ecosystem participation assessment and suggests that, in the examined Factory-X case, governance immaturity, liability allocation and compliance overhead weigh more heavily than technological deficits. For practitioners in machinery and plant engineering, the Scoreboard and Feasibility Gate provide a pragmatic way to pre-assess the economic rationality of Factory-X participation, identify role-specific bottlenecks, prioritise design efforts across demonstrators and roles and structure internal discussions between business, IT and legal functions. For Use Case 2.4, the results suggest that governance improvements and monetisation-model design are more urgent levers than technical prototyping.

Within the broader TriProGenix framework, the Scoreboard and Feasibility Gate address a specific gap by providing an ecosystem-level, transaction-cost-aware feasibility module that complements established methods for situation analysis, requirements specification, and business-model design. These instruments contribute to the envisaged decision-oriented rule set by rendering role and participation choices in Factory-X-type ecosystems assessable in a structured and auditable manner. Future iterations of TriProGenix will therefore integrate the pyramid and Scoreboard more tightly with process modules and governance triggers, enabling consistent evaluation of product-, service-, and ecosystem-level profitability along the transformation trajectory.

7.3 Limitations

Limitations of this study include the single-use-case focus, the limited number of experts and the qualitative nature of the scoring process. The use of one common driver set across all roles preserves comparability but may reduce diagnostic depth for specific roles. Future work should therefore examine to what extent role-specific variants of the Scoreboard, with adapted drivers or anchors, could increase diagnostic depth without sacrificing cross-role transparency. Further research should also calibrate the indices with quantitative data where available, extend the evaluation to additional Factory-X use cases and other federated industrial data ecosystems, and link the qualitative feasibility signals more tightly to scenario-based cash-flow models once pricing schemes and contractual templates mature. The current feasibility results should not be transferred directly beyond Factory-X or to more mature ecosystems, because governance overhead, contractual clarity, and cost allocation may change substantially with ecosystem maturity.

Appendix

A. Overview of the content requirements

Code	Requirement	Chapter
CR 1	Method respects Value Proposition, Value Creation, Value Capture, Value Network and dependencies.	3.1 Business Model
CR 2	Represent hybrid data-, product- & service-based revenue logics	3.2 Servitization
CR 3	Capture lifecycle-wide cost & benefit drivers of servitized offerings	3.2 Servitization
CR 4	Incorporate ex-ante & ex-post transaction-cost components for federated collaboration	3.3. Data Ecosystem
CR 5	Support role-specific parameterisation as functional capabilities	3.3. Data Ecosystem
CR 6	Consider network-level co-creation effects	3.3. Data Ecosystem
CR 7	Reflect data-sovereignty & governance constraints that influence monetary value	3.3. Data Ecosystem

B. Analysis of requirement fulfilment of existing evaluation approaches [68–91]

		Author (Year)	CR1	CR2	CR3	CR4	CR5	CR6	CR7
Explicit	Qualitative	Gassmann et al. (2021b)	●	●	○	○	○	◐	○
		Wirtz (2020)	●	◐	○	○	◐	◐	○
		Zollenkop (2006)	●	◐	◐	◐	○	◐	○
		Urmetzer et al. (2016)	●	◐	○	○	●	●	○
		Allee (2008)	●	◐	○	○	◐	●	◐
	Semi-quantitative	Ahsen et al. (2009)	●	◐	○	○	◐	○	●
		Batocchio et al. (2017)	◐	●	◐	○	●	○	●
		Schallmo (2018)	●	◐	◐	○	◐	○	●
		Mateu and Escribá-Esteve (2019)	○	◐	◐	◐	●	○	●
		Gausemeier et al. (2017)	◐	◐	◐	○	○	◐	○
		Obermaier et al. (2019)	◐	◐	◐	○	◐	○	○
		Liebrecht (2020)	◐	●	●	○	○	◐	○
		Kaufmann and Servatius (2020)	◐	●	○	◐	◐	●	◐
		Pokorni et al. (2021)	○	○	◐	○	◐	○	●
		Tellkamp (2005)	◐	○	●	◐	○	○	○
		Feige et al. (2019)	●	◐	●	◐	◐	◐	◐
		Linde et al. (2021)	●	◐	◐	○	◐	◐	○
		Taschner (2017)	○	○	○	○	○	●	●
		Meierhofer et al. (2022)	◐	◐	◐	○	◐	◐	◐
		Anandarajan and Wen (1999)	◐	○	◐	◐	○	◐	◐
		Brugger (2005)	○	◐	●	○	○	◐	●
		Bensberg et al. (2022)	●	●	●	○	◐	◐	◐
		Gordijn and Akkermans (2001)	●	●	◐	○	●	◐	○
		Eisape (2020)	●	●	◐	◐	◐	●	◐

C. Feasibility Gate results overview

Role	n	B	TC	R	B ≥ TC	R ≥ 0.50	Gate Outcome
Machine Builder	3	0.404	0.867	0.304	No	No	Not Feasible
Component Supplier	3	0.299	1.0000	0.203	No	No	Not Feasible
Factory Operator	3	0.368	0.839	0.355	No	No	Not Feasible
Software Provider	1	0.667	0.933	0.462	No	No	Not Feasible

Figure 4: Feasibility Gate results overview.

D. Scoreboards used during Interview Sessions for Use Case CMLS

Dimension	Code	Driver	Guideline 1	2	3	4	5
Ecosystem & Network (Who)	N-1	Role Coverage & Partner Identification	For UC 2.4, are the roles (provider, consumer, enabler, and federator function) named and committed to (contacts + MoU/LoI), with the contribution and evidence assigned to each role?				
			- No role map - Only generic actor categories listed	- Roles mapped at category level - A few candidate firms named, no outreach recorded.	- All required roles have named candidates - Initial contact made and interest expressed (emails/calls).	- Named partners qualified and aligned on scope - MoU/LoI or draft contracts for a pilot exist.	- Roles contracted (or equivalent binding commitment) - backup options identified for each critical role.
	N-2	Network-effect Potential	To what extent do boundary resources (APIs/SDKs, open standards, partner programs) and reusable additions contribute to the benefits growing as new participants join—as evidenced by the reuse of the same addition across roles/locations?				
			- No observable network effects - participation remains isolated.	- Anecdotal evidence of cross-partner interaction - weak impact on value creation.	- Documented cases of repeatable network effects (e.g., data re-use, role interaction) in limited scope.	- Consistent and systematic network effects across multiple roles - externally visible adoption patterns.	- Strong network effects embedded in ecosystem design - validated across use cases; ecosystem recognized for its network-driven value creation.

			Guideline				
Dimension	Code	Driver	1	2	3	4	5
Value Proposition (What)	W-1	Efficiency Gains	**Does UC 2.4 have consistent evidence of improved process efficiency in the target process that is accepted by the partners involved?**				
			- No observable reductions - Only theoretical potential discussed internally	- Minor improvements claimed - but evidence is anecdotal or not systematically measured	- Structured improvements documented in a limited scope (e.g., one line, one process) - internal reports available.	- Multiple processes or lines show consistent improvements - baseline and results are externally auditable (e.g., certified reports, external benchmark)	- Efficiency improvements embedded in organizational strategy - systematically measured and validated across sites - evidence accepted by ecosystem partners
	W-2	Sustainability & Compliance Gains	**Does UC 2.4 deliver verifiable sustainability/compliance results that comply with the required frameworks and are accepted by customers or auditors?**				
			- No compliance or sustainability measures implemented - no reporting	- Ad-hoc compliance measures (e.g., basic legal requirements) - reporting minimal, limited business relevance.	Structured sustainability processes in place (e.g., partial lifecycle reporting, CO_2 footprints for selected products) - internally documented - moderate stakeholder attention	- Advanced sustainability and compliance approach (e.g., eco-design, supply-chain reporting) - externally auditable reports - customers recognize and sometimes reward these efforts	- Comprehensive sustainability strategy covering full supply chain - externally validated - sustainability contributes visibly to competitive positioning and financial outcomes
	W-3	Insight-driven Revenue Enablement	**To what extent does UC 2.4 deliver repeatable, measurable, attributable revenue effects (e.g., conversion/attach uplift, etc.) that stakeholders attribute to the use case?**				
			- No insight-to-action mechanism in the offering - insights are ad-hoc, not consumable by customers - no reusable artifacts.	- Isolated enablement (decks/demos), but no standard path from signal to offer - manual hand-offs - low repeatability	- Documented mechanism in limited scope (one product/segment): reusable playbooks, value calculators, or embedded analytics - internal use occasional - evidence exists	- Systematic mechanism across multiple offerings: embedded analytics/copilots in product - standard triggers integrated with CRM/CPQ - externally auditable collateral - partners acknowledge reliability.	- Insight-to-action is productized and core to the value proposition portfolio-wide - enablement KPIs (e.g., conversion/attach proxies) are tracked and improved - mechanism is externally recognized and accepted by ecosystem partners.

			Guideline				
Dimension	Code	Driver	1	2	3	4	5
Value Creation (How)	H-1	Data-pipeline CAPEX / OPEX	**To what extent are the CAPEX/OPEX of the data pipeline managed in a structured, evidence-based manner, enabling repeatable, savings across all projects?**				
			- No structured approach to managing CAPEX/OPEX in data pipelines - costs unclear.	- Anecdotal cost savings claimed (e.g., retrofits) - not systematically tracked.	- Documented cost comparisons exist for limited scope (e.g., one asset type) - internal reporting available	- Multiple projects show consistent savings - results externally auditable (e.g., vendor benchmarks, audits)	- Cost management embedded in strategy - savings systematically validated across projects - ecosystem partners accept evidence
	H-2	Analytics & Lifecycle Effort	**To what extent is the analytics lifecycle (collection → validation → monitoring → retraining) managed and logged in such a way that the time and costs involved in the various phases for UC 2.4 can be estimated?**				
			- No structured analytics lifecycle management - models built ad hoc.	- Individual projects track lifecycle costs or effort - but evidence remains anecdotal.	- Documented lifecycle management processes in limited scope - partial reporting available.	- Systematic lifecycle management across multiple analytics projects - externally auditable or benchmarked.	- Lifecycle cost and effort management embedded in strategy - practices validated across organization; ecosystem partners recognize robustness.
	H-3	Inter-disciplinary Coordination Complexity	**To what extent are interdisciplinary coordination mechanisms—such as clear interface responsibilities, change control, and escalation paths—defined and used?**				
			- No structured coordination between disciplines - rework common but undocumented.	- Ad-hoc coordination processes - anecdotal evidence of rework or delays.	- Documented coordination practices for selected projects - partial reporting.	- Systematic coordination mechanisms in place across multiple projects - externally auditable or benchmarked.	- Coordination excellence embedded in organizational routines - validated across business units; recognized externally as best practice.
	H-4	Organisational Readiness & GTM Capability	**To what extent is the market launch of UC 2.4 (catalog entry, SLAs, sales/customer service support) documented, adopted, and regularly reviewed (internally/externally), with improvement measures being implemented?**				
			- No structured readiness for data-driven GTM - no training or capabilities documented.	- Early training or pilot programs exist, - but not systematic - anecdotal evidence of readiness.	- Documented GTM capabilities in limited scope - partial adoption across teams.	- Systematic readiness programs embedded - externally auditable (e.g., certifications, maturity models).	- Readiness and GTM capability embedded in business model - validated across organization - externally recognized as a differentiator.
	H-5	Interoperability Maturity	**How repeatable and auditable is cross-company exchange for UC 2.4—via connectors with issued identities, registered offers/self-descriptions, enforcement of runtime policies, logging, and shared semantics?**				
			- Ad-hoc exchange (files/APIs) without shared semantics, identity, or catalog - no repeatable path.	- Basic syntax-level APIs with bilateral keys - no catalog - semantics ad-hoc - minimal logs.	- Connector-based exchange with issued identities and registered offerings - limited semantic mapping; pilot-repeatable.	- Shared semantic models in use - policies enforced at runtime - multiple partners onboarded through a catalog/broker	- Federation-grade maturity: end-to-end, multi-partner, cross-use-case exchanges - with conformance evidence and operational SLAs

			Guideline				
Dimension	**Code**	**Driver**	**1**	**2**	**3**	**4**	**5**
Value Capture (Value)	V-1	Direct Monetisation Flexibility	**How flexible is monetization for UC 2.4—can at least one accepted logic (subscription, pay-per-use, etc.) with defined base values, measurement, and a verifiable path be billed, and can models be changed later without renegotiating the standard?**				
			- Only one pricing model applied - no flexibility beyond basic product sales.	- Early experiments with service pricing (e.g., subscription or pay-per-use) - but not systematic.	- Multiple monetisation logics used in parallel - documented but applied selectively.	- Monetisation mix applied consistently across offerings - internal processes enable recurring revenue collection.	- Monetisation flexibility embedded in business model - customers accept diverse models; approach externally recognized as a differentiator.
	V-2	Net Ecosystem Rents	**How transparent, predictable, and fair are the ecosystem fees for UC 2.4—is there a documented fee schedule and evidence that participants recognize a positive cost-benefit ratio?**				
			- Fee structures not transparent - participants perceive platform costs as unpredictable or excessive.	- Basic fees or take-rates known but vary case by case - limited transparency.	- Transparent fee schedules available - platform rents broadly accepted but not optimized.	- Platform rents clearly benchmarked and predictable - participants see positive cost–benefit balance.	- Platform rents optimized and widely perceived as fair - pricing structure externally validated and attracts broad participation.
	V-3	Indirect Revenue Uplift	**How clearly are UC 2.4's indirect offers packaged and assigned (catalog entry, authorization, transfer, named owners, documented sales movement), and do partners recognize that they are driving renewals, cross-selling, or product enforcement?**				
			- No observable link between ecosystem participation and indirect revenue.	- Anecdotal cases of upsell/cross-sell through the ecosystem - but not structured or repeatable.	- Some systematic opportunities documented (e.g., add-on services) - moderate adoption internally.	- Indirect revenue streams actively managed - partners recognize recurring benefits.	- Indirect revenue uplift programmatic - embedded in ecosystem strategy - consistently realized across accounts and externally acknowledged.
	V-4	Market-entry Discounts & Incentives	**How clearly defined and institutionalized are the incentives for introducing UC 2.4 (authorization, duration, ownership/approvals, pilot-to-contract plan) so that they measurably lower the barriers to entry and accelerate implementation?**				
			- No incentives - high upfront adoption barriers - negligible uptake of digital offerings.	- Ad-hoc or one-off discounts - limited effect on adoption rates.	- Structured incentives (rebates, bundle discounts, or fiscal support) reduce entry costs and stimulate moderate adoption (+10–20%).	- Comprehensive incentive schemes (rebates, policy support, prototyping subsidies) accelerate adoption by 30–50% across adopter segments.	- Fully institutionalized adoption support mechanisms - with measurable acceleration-with sustainable integration into long-term diffusion strategies.

			Guideline				
Dimension	Code	Driver	1	2	3	4	5
Governance / TCE	G-1	Ex-ante Transaction Costs	**To what extent are the ex-ante transaction costs of UC 2.4 standardized, clearly assigned to specific responsibilities, and recorded using reliable key figures—with evidence of predictability and reduced friction losses?**				
			- No structured handling of pre-contractual costs - effort undocumented.	- Anecdotal reporting of pre-contractual effort - not systematically measured.	- Documented pre-contractual costs for selected projects - internal reports available.	- Systematic cost tracking across multiple projects - results externally auditable or benchmarked.	- Ex-ante transaction cost management embedded in contracting strategy - validated across projects - recognized externally as efficient.
	G-2	Ex-post Transaction Costs	**To what extent are the ex-post transaction costs of UC 2.4 anchored as an operational workflow with roles/tools/KPIs and documented by reports/audits—with evidence of predictability and decreasing rework and disputes?**				
			- No structured handling of post-contractual costs - disputes and rework undocumented.	- Anecdotal evidence of disputes or costs - not systematically measured.	- Documented post-contractual costs for limited scope - partial reporting available.	- Systematic tracking across multiple projects - externally auditable or benchmarked.	- Ex-post transaction cost management embedded in organizational practice - validated across projects - externally recognized as best practice.
	G-3	Sovereignty & Compliance Overhead	**To what extent are sovereignty controls provided as reusable, automated services with designated owners, runbooks, and audit trails—and proven in cross-company exchange processes, so that partners consider the operational effort to be predictable and acceptable?**				
			- No structured sovereignty or compliance measures - overhead unmanaged	- Anecdotal evidence of sovereignty-related overhead (e.g., IAM integration) - not systematically measured.	- Documented overhead for limited scope (e.g., selected compliance projects) - internal reports available	- Systematic overhead tracking across projects - externally auditable or benchmarked.	- Sovereignty and compliance overhead systematically managed - embedded in governance strategy - validated across ecosystem.

References

[1] Paiola, M., Gebauer, H., 2020. Internet of things technologies, digital servitization and business model innovation in BtoB manufacturing firms. Industrial Marketing Management 89, 245–264.

[2] Soellner, S., Helm, R., Klee, P., Endres, H., 2024. Industrial service innovation: Exploring the transformation process to digital servitization in industrial goods companies. Industrial Marketing Management (117), 288–303.

[3] Hoffmann, F., Koch, T., Weber, M., Weigold, M., Metternich, J., 2023. Development of Data-based Business Models to Incentivise Sustainability in Industrial Production. Hannover : publish-Ing.

[4] Hoffmann, F., Weber, M., Weigold, M., Metternich, J., 2022. Developing GAIA-X Business Models for Production. Hannover : publish-Ing.

[5] La Calle, A. de, Freije, I., Larrinaga, M.A., 2025. Digital service innovation in SMEs: identifying the barriers for business model evolution. Journal of Enterprise Information Management.

[6] Hallegatte, S., Shah, A., Lempert, R., Brown, C., Gill, S., 2012. Investment decision-making under deep uncertainty-application to climate change. Policy research working paper.

[7] Otto, B., Mohr, N., Roggendorf, M., Guggenberger, T., 2020. Data sharing in industrial ecosystems: Driving value across entire production lines. McKinsey & Company; Fraunhofer ISST. https://www.mckinsey.de/~/media/mckinsey/locations/europe%20and%20middle%20east/deutschland/publikationen/data%20sharing%20in%20industrial%20ecosystems/mckinsey_article_data_sharing_in_industrial_ecosystems.pdf. Accessed 17 November 2025.

[8] European Commission, 2022. Data Act: Commission proposes measures for a fair and innovative data economy. Press release IP/22/1113. https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_22_1113/IP_22_1113_EN.pdf. Accessed 17 November 2025.

[9] Lohmöller, J., Pennekamp, J., Matzutt, R., Schneider, C.V., Vlad, E., Trautwein, C., Wehrle, K., 2024. The unresolved need for dependable guarantees on security, sovereignty, and trust in data ecosystems. Data & Knowledge Engineering 151, 102301.

[10] Sullivan, C.S., Gemtou, M., Anastasiou, E., Fountas, S., 2024. Building trust: A systematic review of the drivers and barriers of agricultural data sharing. Smart Agricultural Technology 8, 100477.

[11] European Parliament, Council of the European Union, 2023. Regulation (EU) 2023/2854 on harmonised rules on fair access to and use of data (Data Act): Data Act.

[12] Lechowski, G., Krzywdzinski, M., 2022. Emerging positions of German firms in the industrial internet of things: A global technological ecosystem perspective. Global Networks 22 (4), 666–683.

[13] Tardieu, H., 2022. Role of Gaia-X in the European Data Space Ecosystem, in: Otto, B., Hompel, M. ten, Wrobel, S. (Eds.), Designing Data Spaces. Springer International Publishing, Cham, pp. 41–59.

[14] BMWK, 2025. Manufacturing-X – Funding scheme for a competitive, resilient and sustainable industry. BMWK. https://www.bundeswirtschaftsministerium.de/Redaktion/EN/Publikationen/Digitale-Welt/manufacturing-x-program.pdf?__blob=publicationFile&v=6.

[15] Factory-X Consortium, 2025. FACTORY-X – The digital ecosystem. Open Industry 4.0 Alliance Implementation GmbH. https://factory-x.org/. Accessed 26 October 2025.

[16] Peric, M., Durkin, J., Vitezic, V., 2017. The Constructs of a Business Model Redefined: A Half-Century Journey. Sage Open 7 (3).

[17] Fielt, E., 2013. Conceptualising business models: Definitions, frameworks and classifications. Journal of Business Models 1 (1), 85–105.

[18] Osterwalder, A., Pigneur, Y., Tucci, C.L., 2005. Clarifying Business Models: Origins, Present, and Future of the Concept. CAIS 16.

[19] Teece, D.J., 2010. Business Models, Business Strategy and Innovation. Long Range Planning 43 (2), 172–194.

[20] Gassmann, O., Frankenberger, K., Choudury, M., 2020. Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler business model navigator. Carl Hanser Verlag GmbH Co KG.

[21] Osterwalder, A., Pigneur, Y., 2010. Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, 1 online resource (278 s.).

[22] Baines, T.S., Lightfoot, H.W., Benedettini, O., Kay, J.M., 2009. The servitization of manufacturing: A review of literature and reflection on future challenges. JMTM 20 (5), 547–567.

[23] Vandermerwe, S., Rada, J., 1988. Servitization of business: Adding value by adding services. European Management Journal 6 (4), 314–324.

[24] Vendrell-Herrero, F., Parry, G., Bustinza, O.F., O'Regan, N., 2014. Servitization as a driver for organizational change. Strategic Change 23 (5-6), 279–285.

[25] Reim, W., Parida, V., Örtqvist, D., 2015. Product–Service Systems (PSS) business models and tactics – a systematic literature review. Journal of Cleaner Production 97, 61–75.

[26] Riesener, M., Kuhn, M., Boβmann, C., Schuh, G., 2023. Concept for the Portfolio Management of Industrial Solution Providers in Machinery and Plant Engineering. Procedia CIRP 119, 1152–1157.

[27] Tukker, A., 2004. Eight types of product–service system: eight ways to sustainability? Experiences from SusProNet. Bus Strat Env 13 (4), 246–260.

[28] Vargo, S.L., Lusch, R.F., 2004. Evolving to a New Dominant Logic for Marketing. Journal of Marketing 68 (1), 1–17.

[29] Oliva, R., Kallenberg, R., 2003. Managing the transition from products to services. International Journal of Service Industry Management 14, 160–172.

[30] Kohtamäki, M., Rabetino, R., Parida, V., Sjödin, D., Henneberg, S., 2022. Managing digital servitization toward smart solutions: Framing the connections between technologies, business models, and ecosystems 105.

[31] Sjödin, D., Parida, V., Jovanovic, M., Visnjic, I., 2020. Value creation and value capture alignment in business model innovation: A process view on outcome-based business models. J of Product Innov Manag 37 (2), 158–183.

[32] Chesbrough, H., 2002. The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. Industrial and Corporate Change 11 (3), 529–555.

[33] Otto, B., Burmann, A., 2021. Europäische Dateninfrastrukturen. Informatik Spektrum 44 (4), 283–291.

[34] Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a Theory of Ecosystems. SSRN Journal.

[35] Oliveira, M.I.S., Lóscio, B.F., 2018. What is a Data Ecosystem? Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, 1–9.

[36] Schroeder, R., 2016. Big data business models: Challenges and opportunities. Cogent Social Sciences 2 (1), 1166924.

[37] Strnadl, C.F., 2023. A Formal Transaction Cost-Based Analysis of the Economic Feasibility of Ecosystems. arXiv preprint arXiv:2310.03157.

[38] Gaia-X Association for Data and Cloud, 2022. Gaia-X Architecture Document – 22.04 Release. https://gaia-x.eu/wp-content/uploads/2022/06/Gaia-x-Architecture-Document-22.04-Release.pdf.

[39] Otto, B., Seidelmann, J., Schmelting, J., Sauer, O., 2023. Manufacturing-X Data Space Study.

[40] Foss, N.J., Saebi, T., 2017. Fifteen years of research on business model innovation: How far have we come, and where should we go? Journal of Management 43 (1), 200–227.

[41] Geissdoerfer, M., Vladimirova, D., van Fossen, K., Evans, S., 2018. Product, service, and business model innovation: A discussion. Procedia Manufacturing 21, 165–172.

[42] Parida, V., Sjödin, D., Reim, W., 2019. Reviewing Literature on Digitalization, Business Model Innovation, and Sustainable Industry: Past Achievements and Future Promises. Sustainability 11 (2), 1–18.

[43] Keeney, R.L., Raiffa, H., 1993. Decisions with multiple objectives: preferences and value trade-offs. Cambridge University Press.

[44] Saaty, R.W., 1987. The analytic hierarchy process—what it is and how it is used. Mathematical Modelling 9 (3-5), 161–176.

[45] Al-Debei, M.M., Avison, D., 2010. Developing a unified framework of the business model concept. EUROPEAN JOURNAL OF INFORMATION SYSTEMS 19 (3), 359–376.

[46] Kett, H.J., Evcenko, D., Falkner, J., 2023. Equipment-as-a-Service (EaaS).

[47] Rabetino, R., Harmsen, W., Kohtamäki, M., Sihvonen, J., 2018. Structuring servitization-related research. IJOPM 38 (2), 350–371.

[48] Zheng, P., Wang, Z., Chen, C.-H., Pheng Khoo, L., 2019. A survey of smart product-service systems: Key aspects, challenges and future perspectives. ADVANCED ENGINEERING INFORMATICS 42, 100973.

[49] Wirth, J., Erb, S., Hoffmann, F., Metternich, J., Bauernhansl, T., 2024. Framework for the Introduction of Data-driven and Service-oriented Business Models in Mechanical and Plant Engineering Companies. Procedia CIRP 130, 1124–1129.

[50] Adner, R., 2017. Ecosystem as Structure: An Actionable Construct for Strategy. Journal of Management 43 (1), 39–58.

[51] Moore, J., 1993. Predators and Prey: A New Ecology of Competition. Harvard business review 71, 75–86.

[52] Corrado, C., Haskel, J., Iommi, M., Jona Lasinio, C.S., 2022. The value of data in digital-based business models: Measurement and economic policy implications. OECD Economics Department Working Papers 1723.

[53] Hoffmann, F., Lang, E., Metternich, J., 2023. Seizing the value of data: Selecting appropriate pricing strategies for data-based services in manufacturing. Procedia CIRP 120, 159–164.

[54] Laney, D.B., 2017. Infonomics: how to monetize, manage, and measure information as an asset for competitive advantage. Routledge.

[55] Hoffmann, F.S., 2024. Methode zur Bewertung unternehmerischer Erfolgsaussichten in der Entstehungsphase eines datenbasierten Geschäftsmodells im Kontext von Industrie 4.0. Dissertation. Shaker, Darmstadt.

[56] Kolagar, M., Parida, V., Sjödin, D., 2022. Ecosystem transformation for digital servitization: A systematic review, integrative framework, and future research agenda. Journal of Business Research 146, 176–200.

[57] Hevner, A., 2007. A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19 (2), 87–92.

[58] Hevner, A., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research. MIS Quarterly 28 (1), 75–105.

[59] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–77.

[60] vom Brocke, J., Riemer, K., Niehaves, B., Plattfaut, R., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. ECIS 2009 Proceedings (161), 1–15.

[61] Webster, J., Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), xiii–xxiii.

[62] Wohlin, C., 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering, in: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, New York, NY, USA. ACM, New York, NY, USA, pp. 1–10.

[63] Albers, A., Burkhardt, N., Meboldt, M., Saak, M., 2005. SPALTEN Problem Solving Methodology in the Product Development, in: Engineering Design and the Global Economy: 15th International Conference on Engineering Design - ICED 2005, Melbourne, 15. - 18. August 2005. Hrsg.: A. Samuel. The Design Society, p. 90671.

[64] Wirth, J., Hermann, M., Bauernhansl, T., 2025. Navigating towards New Businesses: An Implementation Framework for Data-Driven and Service-Oriented Business Models. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[65] Wirth, J., Neller, D., Bauernhansl, T., 2025. The Servitization Ecosystem Canvas (SEC): Towards a framework for developing service-oriented and data-driven business models in ecosystems. Procedia CIRP 134, 360–365.

[66] Evans, D.S., Schmalensee, R., 2016. Matchmakers: The new economics of multisided platforms. Harvard Business Review Press.

[67] Rochet, J.-C., Tirole, J., 2003. Platform competition in two-sided markets. Journal of the european economic association 1 (4), 990–1029.

[68] Ahsen, A. von, Kuchenbuch, A., Heesen, M., 2009. Leitfaden: Bewertung von Innovationen im Mittelstand, in: Ahsen, A. von (Ed.), Bewertung von Innovationen im Mittelstand (German Edition). Springer, Dordrecht, pp. 39–74.

[69] Allee, V., 2008. Value network analysis and value conversion of tangible and intangible assets. Journal of Intellectual Capital 9 (1), 5–24.

[70] Anandarajan, A., Wen, H.J., 1999. Evaluation of information technology investment. Management Decision 37 (4), 329–339.

[71] Batocchio, A., Ferraz Minatogawa, V.L., Anholon, R., 2017. Proposal for a Method for Business Model Performance Assessment: Toward an Experimentation Tool for Business Model Innovation. Journal of Technology Management & Innovation 12 (1), 61–70.

[72] Bensberg, F., Czarnecki, C., Grob, H.L., 2022. Wirtschaftlichkeitsbewertung von Smart Services mit vollständigen Finanzplänen. Smart Services, 225–251.

[73] Brugger, R., 2005. Der IT Business Case. Kosten erfassen und analysieren - Nutzen erkennen und quantifizieren - Wirtschaftlichkeit nachweisen und realisieren., 2nd ed. Springer, Berlin.

[74] Eisape, D., 2020. Comparing Platform Business Models. Xx-xx Pages / Nordic Journal of Media Management, Vol. 1 No. 3 (2020).

[75] Feige, B.A., Fabry, H., Gudergan, G., Krechting, D., 2019. Aufwandsbemessung zur Umsetzung neuer Geschäftsmodelle auf der Basis von Distanzmatrizen, in: Stich, V., Schumann, J.H., Beverungen, D., Gudergan, G., Jussen, P. (Eds.), Digitale Dienstleistungsinnovationen. Smart Services agil und kundenorientiert entwickeln. Springer Vieweg, Berlin, Heidelberg, pp. 147–165.

[76] Gassmann, O., Frankenberger, K., Choudury, M., 2021. Geschäftsmodelle entwickeln: 55+ innovative Konzepte mit dem St. Galler Business Model Navigator, 3., überarbeitete und erweiterte Auflage ed. Hanser, München, 451 pp.

[77] Gausemeier, J., Wieseke, J., Echterhoff, B., Isenberg, L., Koldewey, C., Mittag, T., Schneider, M., 2017. Mit Industrie 4.0 zum Unternehmenserfolg. Integrative Planung von Geschäftsmodellen und Wertschöpfungssystemen., 110 pp.

[78] Gordijn, J., Akkermans, H., 2001. e 3-value : Design and Evaluation of e-Business Models, in: .

[79] Kaufmann, T., Servatius, H.-G., 2020. Das Internet der Dinge und Künstliche Intelligenz als Game Changer: Wege zu einem Management 4.0 und einer digitalen Architektur. Springer Vieweg, Wiesbaden, Heidelberg, 211 pp.

[80] Liebrecht, C., 2020. Entscheidungsunterstützung für den Industrie 4.0-Methodeneinsatz – Strukturierung, Bewertung und Ableitung von Implementierungsreihenfolgen.

[81] Linde, L., Sjödin, D., Parida, V., Gebauer, H., 2021. Evaluation of Digital Business Model Opportunities. Research-Technology Management 64 (1), 43–53.

[82] Mateu, J.M., Escribá-Esteve, A., 2019. Ex-Ante Business Model Evaluation Methods. 25-47 Pages / Journal of Business Models, Vol 7 No 5 (2019): Journal of Business Models.

[83] Meierhofer, J., Benedech, R., Schweiger, L., Barbieri, C., Rapaccini, M., 2022. Quantitative Modelling of the Value of Data for Manufacturing SMEs in Smart Service Provision. ITM Web Conf. 41, 4001.

[84] Obermaier, R. (Ed.), 2019. Handbuch Industrie 4.0 und digitale Transformation: Betriebswirtschaftliche, technische und rechtliche Herausforderungen. Springer Gabler, Wiesbaden, Heidelberg, 779 pp.

[85] Pokorni, B., Braun, M., Knecht, C., 2021. Menschzentrierte KI-Anwendungen in der Produktion. Praxiserfahrungen und Leitfadenzu betrieblichen Einführungsstrategien.

[86] Schallmo, D., 2013. Geschäftsmodelle erfolgreich entwickeln und implementieren: Mit Aufgaben und Kontrollfragen; [Lehrbuch. Springer Gabler, Berlin, Heidelberg, 270 pp.

[87] Taschner, A., 2017. Business Cases: Ein anwendungsorientierter Leitfaden, 3. Aufl. 2017 ed. Springer Gabler, Wiesbaden.

[88] Tellkamp, C., 2005. Finanzielle Bewertung von Ubiquitous-Computing-Anwendungen, in: Fleisch, E., Mattern, F. (Eds.), Das Internet der Dinge. Ubiquitous Computing und RFID in der Praxis: Visionen, Technologien, Anwendungen, Handlungsanleitungen. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, pp. 315–327.

[89] Urmetzer, F., Martinez, V., Neely, A., 2016. The Ecosystem Value Framework:Supporting Managers to UnderstandValue Exchange between CoreBusinesses in Service Ecosystems.

[90] Wirtz, B.W., 2020. Business Model Management: Design - Process - Instruments, 2nd ed. 2020 ed. Springer International Publishing; Imprint Springer, Cham, 149 pp.

[91] Zollenkop, M. (Ed.), 2006. Geschäftsmodellinnovation: Initiierung eines systematischen Innovationsmanagements für Geschäftsmodelle auf Basis lebenszyklusorientierter Frühaufklärung. Gabler Verlag / GWV Fachverlage GmbH Wiesbaden, Wiesbaden.

Biography

Jonas Wirth (*1993) has been a doctoral candidate at Fraunhofer IPA/ IFF, of the University of Stuttgart in cooperation with TRUMPF SE + Co. KG since 2022. His research interest is the scale-up of service-oriented business models in value networks. In the research project Factory-X, Mr. Wirth is responsible for the business model project.

Fabian Meldt (*2001) is a master's student of Business Administration and Mechanical Engineering at Darmstadt University of Technology and has conducted research in collaboration with TRUMPF SE + Co. KG on the evaluation of data driven business models and ecosystems.

Felix Hoffmann (*1993) has been a researcher at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2019. His research interest are data-driven Business Models in manufacturing. From 2022 to 2025 Mr. Hoffmann was head of research at the Production Management department of PTW and since 2025 he is COO at autonomIQ Technologies GmbH.

Matthias Weigold (*1977) is the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology with responsibility of the manufacturing and energy efficiency research groups since 2019. Prof. Weigold contributes many years of industrial experience from management positions at Heidelberger Druckmaschinen AG and SAP SE.

Thomas Bauernhansl (*1969) has been Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since September 2011. He is also Director of the Institute for Industrial Manufacturing and Management IFF at the University of Stuttgart. His scientific focus is on digital and biological transformation, production organisation, factory planning, automation and robotics.

8th Conference on Production Systems and Logistics

Design-Framework for Product-Service Systems in Subscription-Based Business Models

Marius Höding[1,4], Jonas Wirth[2,3,5], Sarah Ekanayake[6], Thomas Prefi[1], Thomas Bauernhansl[2,3]

[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen, Aachen, Germany*
[2]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany*
[3]*IFF of University of Stuttgart, Stuttgart, Germany*
[4]*Accenture, Kronberg im Taunus, Germany*
[5]*TRUMPF, Ditzingen, Germany*
[6]*IPEK of Eastern Switzerland University of Applied Sciences, Rapperswil, Switzerland*

Abstract

The increasing connectivity enabled by Industry 4.0 is creating new opportunities for business models in the machinery and plant engineering sector. In particular, as-a-service concepts, such as subscription or usage-based models, promise advantages over traditional transactional business models: recurring revenue streams and stronger customer relationships for manufacturers, as well as greater flexibility and lower capital expenditure for users. However, the full potential of these models can only be realized if machine tools are designed as Product-Service Systems (PSS) specifically tailored to service-oriented usage. To date, systematic approaches that translate requirements of service-oriented business models into concrete design and architectural specifications for both hardware and digital infrastructure are lacking. Key considerations such as usage metering, modular and lifetime-optimized hardware, secure data architecture and the integration of sensor systems, edge computing and platform services remain insufficiently addressed. This paper proposes a methodological framework that bridges the gap between business model logic and technical implementation. It encompasses the implications of service-oriented business models on the physical and digital design elements of machine tools. By providing a structured framework, the paper aims to guide the development of sustainable, service-oriented PSS in the machine tool industry.

Keywords

Business Model Innovation; Data-driven; Servitization; as a service; Product-Service Systems; Production Systems; TriProGenix

1. Motivation

Manufacturers in the mechanical and plant engineering sector face the dual challenge of remaining competitive under volatile market conditions and increasing cost pressures, while meeting customer demands for more individualized and service-enhanced offerings [1,2]. Innovative business models, such as subscription-based or as-a-service concepts, implemented within product-service systems, promise recurring revenue streams and stronger customer relationships for providers, while enabling improved asset utilization and a shift from capital to operational expenditures for users [3,4]. Despite these prospects, service-oriented PSS in the machinery industry have not fully realized their anticipated potential [5,6]. A central reason is the lack of systematic methodologies that translate business model characteristics into corresponding technical and architectural requirements for the underlying systems [5,7–9].

DOI: https://doi.org/10.15488/20998
ISSN: 2701-6277

Existing implementation frameworks, such as TriProGenix developed by Wirth in prior research [10,11], support mechanical and plant engineering firms in adopting service-oriented and data-driven business models at strategic and organizational levels. However, these frameworks primarily address business model implementation and do not provide detailed guidance on how subscription-based business models can be operationalized into technical architecture and design decisions for machine tools conceived as cyber-physical product-service systems (CPPSS). To address this gap, the present paper adopts a technical design perspective and introduces a complementary three-phase method that leverages business model and implementation outputs as inputs to systematically derive architecture requirements and design recommendations for CPPSS in subscription-based business models within the machine tool domain.

2. Fundamentals

2.1 Product-service systems (PSS) and cyber-physical product-service systems (CPPSS)

Traditionally, companies in the mechanical and plant engineering sector have primarily developed, produced and marketed physical products. However, increasing competitive pressure, market volatility and evolving customer expectations are forcing firms to offer more flexible and customized solutions. One promising strategy to address these developments is the implementation of product-service systems (PSS), which integrate physical products with industrial services [12]. Goedkoop et al. define PSS as *"marketable set of products and services that jointly fulfil specific customer needs"* [13].

When mechatronic systems are further equipped with networking, data processing, and communication capabilities that enable adaptive behavior during operation, they are referred to as cyber-physical systems (CPS) according to the German Guideline VDI 2206 [14]. Combining the principles of PSS and CPS leads to cyber-physical product-services systems (CPPSS), which are defined as *"product-service systems equipped with cyber, physical, networking, computing and communication components that offer built-in intelligent capabilities to achieve higher collaboration, interaction, efficiency, usability, appeal and value co-creation for customers and other stakeholders"* [15].

2.2 Servitization and digital servitization

The transformation from product-centric towards solution-oriented value offerings was introduced into the academic discourse already in 1988 under the term servitization [16]. More recent publications similarly highlight the shift in value creation from physical products towards services and usage-oriented offerings. [17,18]. Along this transformation line, manufacturers increasingly incorporate digital components into their physical offerings, thereby creating new hybrid value propositions [19]. The combination of servitization and the ongoing digitalization of products is referred to as digital servitization. [18,20]. Kohtamäki et al. define digital servitization as a *"transition toward smart product-service-software systems that enable value creation and capture through monitoring, control, optimization and autonomous function"* [21].

2.3 Service-oriented business models

In academic literature, different understandings of the term business model exist [22]. A widely used description is provided by Gassman et al., who argue that a business model specifies who the customers are, what value is offered to them, how this value is created and why the offering is profitable for the firm. Modifying, recombining or extending existing business model elements is commonly referred to as business model innovation. [23] Driven by ongoing servitization, value creation increasingly shifts from products toward services, resulting in more intensive and long-term customer relationships. A prominent service-oriented business model, originating mainly from the software industry, is the subscription or "as-a-service" model, which has recently gained relevance in mechanical and plant engineering. In such models, customers receive continuous access to products or services in exchange for recurring payments. [24]

Modern subscription models additionally feature a strong alignment of interest between provider and customer, as both parties benefit from continuous improvements in the performance and productivity delivery by the Product-Service System. Against this background, Liu et al. define subscription business models as: *"Subscription business models in the machinery and equipment industry describe a long-term relationship between suppliers and customers who receive recurring values for recurring payments. Based on consequent customer centricity, all generated values are designed towards the customers' problems and needs, especially to a continuous increase of their productivity. The values are transported not only by physical products, but also by all digital products and services that are required and desired by the customers." [25]*

3. Related Research

To position the paper within the current state of research, a structured literature review is conducted to identify relevant approaches in the context of (CP)PSS and as-a-service models in machinery and plant engineering. The review focuses on contributions that address the design, architecture, or implementation of digital and cyber-physical (CP)PSS.

Based on this analysis, six representative research approaches are described, analyzed and evaluated that provide conceptual or technological foundations relevant to this work. The identification process follows a systematic procedure inspired by Weber and Watson for literature reviews [26]. The following section summarizes and evaluates these approaches with respect to their contribution to the derivation of technical requirements for (CP)PSS and their applicability to the machine tool domain. The evaluation criteria for the identified approaches include the application scenario, the investigated product and service characteristics, and the recommendation for future system design.

The approach from Ebi proposes a four-stage methodology for implementing subscription-based business models in mechanical and plant engineering. The method derives a generic description model of industrial subscription models, identifies implementation activities and their dependencies, and concludes with a process model for the introduction of subscription offerings. [27] While the approach provides a holistic implementation logic for business model transformation, it does not systematically consider the implications on PSS engineering and its technical architecture.

Alaluss et al. propose a six-stage procedure for transforming ownership-based transactional business models into pay-per-use models for complex production systems. The method includes a readiness assessment, current state modelling through Value Proposition Canvas, Business Model Canvas and customer journey mapping, followed by the design of a pay-per use model and a final quantification of wear costs, technical gaps and pricing mechanisms. [28] While this approach provides relevant guidance for business model transformation, it does not address the technical implications for PSS engineering.

Mennenga et al. present an integrated architecture and development approach combining PSS with Cyber-Physical Production Systems (CPPS). The proposed architecture consists of two main systems: (1) The business model logic, represented through the Business Model Canvas, and (2) the CPPS, which itself comprises a physical layer (e.g. data generation through sensors) and a cyber layer (data processing, modelling, analysis and decision making). In addition, the authors introduce a holistic development framework that links business model design and technical architecture. [29] While the approach effectively integrates conceptual and technical dimensions, it lacks the derivation of concrete architectural or design guidelines for implementing service-oriented CPPSS in practice.

Rizvi and Chew propose a design methodology for CPPSS. Building on a initial conceptual definition of CPPSS, the authors derive a meta-model describing the integration of PSS and Cyber Physical Systems (CPS). Based on an analysis of existing CPS and PSS development methods, a CPPSS design method is derived in the form of a relational framework linking phases, activities, stakeholders, artefacts and their

relationships. [15] While the concept provides a coherent methodological foundation for designing CPPSS, it does not consider the specific characteristics of subscription-based business models nor the industrial context of machinery and plant engineering.

Hicking develops a comprehensive method for specifying intelligent products in mechanical engineering, addressing the methodological gap between planning and conceptual design activities in product development. The approach defines six generic development objectives and twelve generic use cases for intelligent products. Subsequently, the author derives digitalization requirements and analyses their interrelations with the identified use cases. The results are consolidated into a three-step design model for intelligent product development: (1) definition of objectives and boundary conditions, (2) solution exploration through the selection of use cases and analysis of digitalization requirements, and (3) prioritization and decision-making on implementation. [30] While the method provides a substantial contribution to the systematic design of intelligent products, it does not derive specific design recommendations for PSS in the machinery and plant engineering domain.

In addition to these approaches, implementation frameworks have emerged that address the organizational adoption of service-oriented and data-driven business models in mechanical and plant engineering. TriProGenix (TPG), developed by the authors in prior research, exemplifies such a framework. It introduces a three-dimensional implementation logic comprising: (i) a development dimension encompassing strategy, resources, organizational structures, ecosystem relationships, and cultural aspects; (ii) a process dimension structuring the transformation into modular processes with defined gates, triggers, and application paths; and (iii) a method dimension linking appropriate analysis, creativity, decision-making, and project management methods. Furthermore, an operationalization framework incorporating input–throughput–output (ITO) logic, KPI and evidence structures, role definitions, and governance mechanisms supports the planning, steering, and monitoring of the transformation. [10,32,31] Nevertheless, TPG primarily addresses strategic, organizational, and ecosystem-related aspects and does not provide detailed technical design principles for CPPSS at the machine tool level.

	Approach is in the area of…			Approach consideres…			Approach …
	… industrial product-service systems	… service-oriented business models	…machine tool industry	…characteristics of servic-oriented business models	… characteristics of CPPSS	… interdependencies between BM and CPPSS	…derivates design recommendations for CPPSS
Ebi	◕	●	●	◕	○	○	○
Alaluss et al.	◕	●	◑	◑	◑	○	◑
Mennega et al.	◕	◔	◕	◔	◕	◔	◔
Rizvi and Chew	◑	○	◔	○	◑	◑	○
Hicking	◕	◑	◕	◔	●	◕	◑
Wirth et al.	◕	●	●	◕	◑	◔	◔

○ Not fullfilled ◔ hardly fullfilled ◑ partially fullfilled ◕ mostly fullfilled ● completely fullfilled

Figure 1: Evaluation of existing approaches

A wide range of approaches and methods for developing products, services, subscription-based business models, PSS and CPPSS can be found in the scientific literature. However, as shown in Figure 1, the analysis of existing work reveals that none of these approaches sufficiently addresses the relationship between the logic of subscription-based business models and the technical design of the corresponding CPPSS. Empirical evidence indicates that the share of companies achieving full servitization continues to fall short of

expectations. This is emphasized by a scientific survey with 95 companies in the machinery sector conducted by Ardolino et al., in which both the “Shift of company emphasis from a product-centric logic to service-centric logic” as well as “acquiring or having the appropriate technological infrastructure available” are identified as main barriers toward servitization. [5,7,6] Therefore the subsequent section outlines a methodology developed by the authors, which aims to close this practical and research gap by systematically linking the characteristics of subscription-based business model with the architectural requirements of CPPSS.

4. Methodology

Building on the identified research gaps, industry needs [5,7] and the prior work of the authors [10,9,31] this section introduces the methodological framework developed in this study. The proposed approach seeks to close the gap between the logic of service-oriented, subscription-based business models and the technical architecture and design of machine tools conceived as CPPSS. The method aims to offer providers a comprehensive guideline for addressing all relevant design criteria for CPPSS, facilitates targeted and efficient time-to-market processes and support the fulfilment of the increased quality requirements inherent to service-based customer engagements. It is intended to complement existing implementation frameworks, such as TriProGenix, by explicitly addressing the technical implications of business model characteristics for CPPSS design.

The proposed methodological framework, summarized in Figure 2, is structured into three phases and aims to translate business model decisions into technical architecture requirements for CPPSS in subscription-based business models.

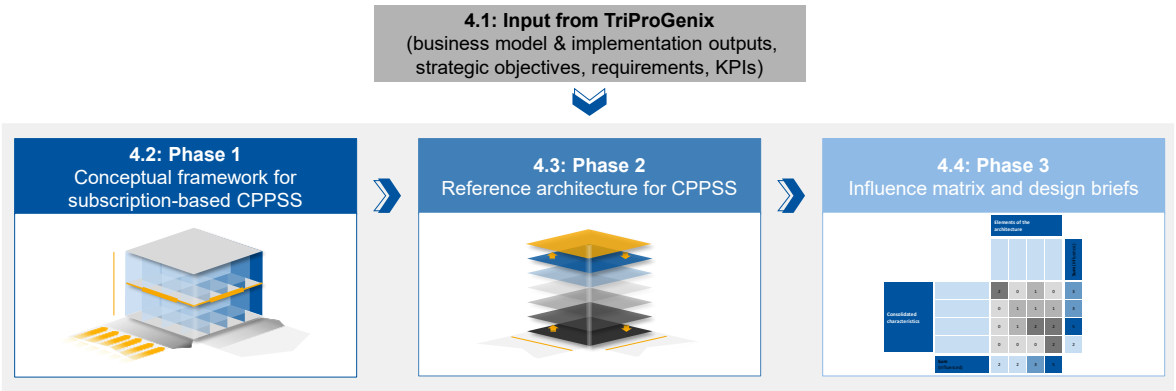


Figure 2: Three-phase methodological framework for deriving architecture requirements for subscription-based CPPSS based on the input of TriProGenix

It builds on outputs from service-oriented business model development and implementation in mechanical and plant engineering, such as business model specifications, strategic objectives, target KPIs, evidence structures, and organizational, process, and ecosystem designs derived from frameworks like TriProGenix. These inputs are used to instantiate and prioritize characteristic categories, populate the conceptual framework, configure the reference architecture, and focus the influence matrix and design briefs on the most critical architecture elements. Figure 2 illustrates this flow, showing how implementation framework outputs serve as inputs and how the method systematically generates architecture requirements and design briefs for CPPSS in the machine tool domain.

4.1 Input: Methodological approach TriProGenix

Implementation frameworks such as TriProGenix (TPG) represent a key element of the knowledge base for the proposed methodological framework. As outlined in Section 3, TPG provides a three-dimensional implementation logic that supports mechanical and plant engineering firms in adopting service-oriented and data-driven business models at strategic and organizational levels (see Appendix 1). Its outputs include a

specification of the service-oriented business model (e.g., target segments, value proposition, revenue logic, contractual setup), strategic and operational objectives, KPIs and evidence requirements, as well as organizational, process, and ecosystem designs with defined roles and governance mechanisms.

In the three-phase method, these outputs serve as structured inputs. In Phase 1, they inform the instantiation and prioritization of characteristic categories and the analysis along the relational, lifecycle-oriented, and perspective dimensions of the conceptual framework. In Phase 2, they guide the configuration of the five-layer reference architecture, for example, by defining which services must be supported at the Service & BM Layer and which KPIs and evidence requirements drive data acquisition and analytics on lower layers. In Phase 3, they support the weighting of influences in the influence matrix and the formulation of design briefs for critical architecture elements. This integration ensures that TPG and the proposed method address complementary levels of digital servitization: organizational implementation on the one hand, and technical architecture and system design on the other.

4.2 Phase 1: Conceptual framework and characteristics of subscription-based PSS

Building on the methodological structure outlined in Section 4, this section introduces the conceptual framework and characteristic categories forming the core of Phase 1. To systematically derive design recommendations for CPPSS in subscription-based business models within mechanical and plant engineering, the framework enables a multi-perspective representation of relevant CPPSS characteristics and provides the foundation for a structured analysis. Based on a comprehensive review of scientific literature on subscription models, CPPSS and related concepts, three main descriptive dimensions can be identified, as shown in Figure 3.

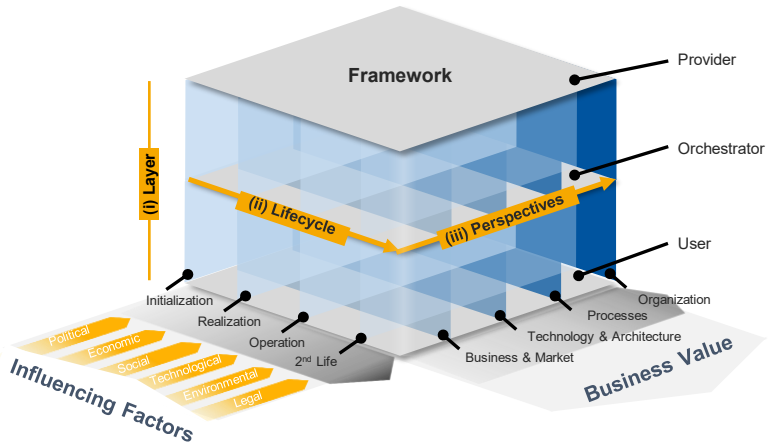


Figure 3: Framework for identifying characteristic features of subscription-based PSS

(i) The relational dimension ("layer") captures the interactions between ecosystem actors at different levels, including providers and developers with their partners (e.g. suppliers or engineering collaborators), customers and users with their respective partner networks, as well as potential intermediaries such as operators or orchestrators. (ii) The "lifecycle"-oriented dimension refers to the chronological stages of a CPPSS, spanning from initiation and conceptualization through development, realization, and operation to potential reuse in a "second life" phase. (iii) The "perspective" dimension reflects the various viewpoints from which a CPPSS can be analyzed, including economic (business objectives, value creation), technological-architectural, organization and processual perspectives. Moreover, external influencing factors are additionally integrated following the PESTEL framework: Political, Economic, Social, Technological, Environmental, Legal [33]. Together, the framework allows for a holistic understanding of CPPSS, where business value emerges from the interplay between these three dimensions. The intersection of the three dimensions within the conceptual framework results in a total of 48 combinations, such as Provider – Operational Phase – Economic Perspective. For each intersection, a characteristic guiding

question can be formulated, for instance: Which descriptive features influence, from the providers perspective, the design of a subscription-based CPPSS during the operation phase to ensure economic viability?

Based on these 48 guiding questions, specific descriptive characteristics of subscription-based CPPSS were derived. A comprehensive literature review initially identified 388 individual characteristics, which were subsequently consolidated and clustered into 36 representative features grouped into 12 characteristic categories. These categories represent the core characteristics of subscription-based CPPSS, as shown in Figure 6 (see Appendix). These feature groups form the systematic basis for deriving technical and architectural requirements for the design of CPPSS within subscription-based business models. The framework can be instantiated using business model specifications, KPI definitions, and organizational information provided by implementation frameworks such as TPG, as outlined in Sections 4.1 and 4.2.

4.3 Phase 2: Derivation of a Reference architecture for CPPSS

To systematically derive design requirements for CPPSS based on the characteristics of subscription-based models, it is necessary to develop a suitable reference architecture. For this purpose, various reference architectures from both literature and industrial practice were analyzed and consolidated into a structure tailored to the objectives of this study (see also [9]). The proposed reference architecture comprises five layers, each representing a distinct functional level within the cyber-physical PSS, as shown in Figure 4.

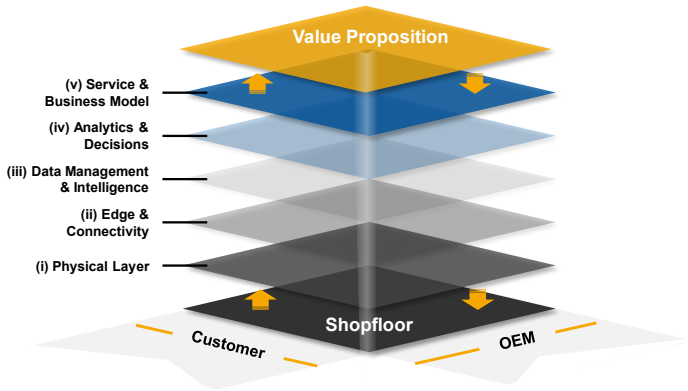


Figure 4: Reference Architecture of CPPSS

- **(i) Physical Layer:** Encompasses the tangible components of the machine tool or PSS, including actuation, mechanics, energy supply and discharge, control units, sensors and other data sources.
- **(ii) Edge & Connectivity Layer:** Contains elements required for local data processing and communication beyond the system boundary, such as edge gateways, middleware or communication technologies.
- **(iii) Data Management & Intelligence Layer:** Integrates data storage, analytics capabilities, event and streaming services as well as identity and access management.
- **(iv) Analysis & Decision Layer:** Comprises monitoring, prediction, prescription, simulation and control functionalities that enable data-driven decision-making.
- **(v) Service & Business Model Layer:** Represents interaction with the customer through dashboarding, billing, contract management and other (digital) services.

Together, these layers form the structural backbone of subscription-based CPPSS, linking physical and digital system components as well as the service proposition to the overarching customer value proposition.

4.4 Phase 3: Analysis of implications and derivation of design recommendations

To derive concrete design recommendations for subscription-based CPPSS, the next methodological step maps the characteristic features identified in Section 4.2 onto the reference architecture defined in Section

4.3 using an influence matrix. This mapping is carried out individually for the specific CPPSS under consideration. First, the characteristic features are reviewed for relevance and completeness, subsequently, the architectural elements are examined to ensure their suitability for the respective system, allowing additional or modified elements to be incorporated where necessary. The adapted features are then linked to the architectural components and the influence is assessed on an ordinal scale from 0 ("no influence") to 2 ("strong influence"). The assessment follows a systematic influence logic structured around three guiding questions: (1) Mechanism: Why does the characteristic exert influence on this architectural element? (2) Direction: In what way does the characteristic alter the element (e.g. more, less, more stable, …). (3) Goal: What business-model-related outcome is the characteristic intended to achieve? Figure 7 (see Appendix) shows the mapping between characteristics and the architecture. Design requirements are subsequently derived at an aggregated level, i.e., for each combination of characteristic group and architectural layer. The derivation follows the logic "Characteristic X → requires capability Y → implies design requirement Z". For example, the characteristic "customer interaction" needs continuous availability of machine data, which leads to the design requirement that "sensor systems must be continuously accessible and serviceable".

Overall, this process generates 60 requirement groups (12 characteristic groups x 5 architectural layers), which together constitute the technical requirements for developing a subscription-enabled CPPSS. Based on these requirement groups, concrete design recommendations can be derived into design briefs (see Appendix 4, Figure 8). The prioritization and selection of design briefs are guided by the objectives, KPIs, and governance requirements defined in the preceding implementation framework.

5. Validation

The methodological approach presented in this paper was developed and initially applied within the research project X-Forge, which explores the implementation of "Everything-as-a-Service" concepts in mechanical and plant engineering [34]. Within this project, the TriProGenix framework provides the strategic and organizational backbone for developing and scaling service-oriented and data-driven business models. Its practical robustness has been demonstrated in four industrial cases at TRUMPF: (1) development of a five-phase onboarding framework for the Condition Monitoring PSS, (2) design and refinement of a service-oriented business model and PSS configuration for the Cutting Assistant, (3) further development of the Flex Cell as a modular automation PSS for bending machines, and (4) retrospective analysis of the Equipment-as-a-Service "Pay-per-Part" business model. Across these cases, TriProGenix enabled structured context analysis, clarification of roles and decision logics, and derivation of actionable recommendations for business model design and organizational alignment. The three-phase method introduced in this paper extends these implementation-oriented results toward technical system design. In an initial industrial application with a machine tool OEM, outputs from TriProGenix, such as a defined subscription-based business model, strategic objectives, target KPIs, and organizational and process setup, served as inputs to instantiate the conceptual framework for subscription-based CPPSS, configure the five-layer reference architecture, and populate the influence matrix. Based on this mapping, design briefs were derived for selected architecture elements, specifying requirements for machine hardware, edge and connectivity infrastructure, data management and analytics capabilities, and customer-facing digital services. The method and resulting design briefs were reviewed with first experts from OEMs. Feedback indicated that the architectural implications align broadly with the existing development roadmap and help make implicit design assumptions explicit and traceable, while highlighting the need for further applications across additional CPPSS cases and for tool support to facilitate adoption by multidisciplinary development teams.

6. Conclusion and Outlook

The approach presented in this paper introduces a structured method for deriving technical and architectural requirements for subscription-based cyber-physical product-service systems (CPPSS) in the machine tool

domain. Building on outputs from the TriProGenix implementation framework, the method comprises three phases: (i) instantiating and prioritizing characteristic features of subscription-based CPPSS using a conceptual framework, (ii) defining and configuring a five-layer reference architecture, and (iii) mapping these characteristics onto the architecture via an influence matrix to derive system-specific design briefs and requirements. This process translates business model decisions into concrete implications for machine hardware, connectivity, data management, analytics, and service-facing components. The main contribution lies in bridging organizational implementation and technical system design, which offers tangible commercial value by enabling customer-centric service design, thereby increasing the likelihood of market success. While TriProGenix supports mechanical and plant engineering firms in designing and implementing service-oriented and data-driven business models at strategic and organizational levels, the proposed method complements this by focusing on the technical level and explicitly linking subscription business model logic to CPPSS architecture elements across physical, edge, data, analysis, and service layers. Development of the method builds on prior empirical work with TriProGenix in multiple industrial cases, and an initial application in the X-Forge project with a machine tool OEM indicates that the derived architectural implications align with existing development roadmaps and help make implicit design assumptions explicit and traceable. Current results represent an initial validation step. Future research should extend empirical assessment across additional CPPSS cases and asset classes, evaluate the method's impact on development outcomes using quantitative and longitudinal criteria, and provide tool support to facilitate adoption by multidisciplinary development teams. Further integration with implementation frameworks such as TriProGenix and with engineering toolchains, e.g., within model-based systems engineering environments, appears particularly promising. Overall, the proposed methodology contributes to a closer integration of business model thinking and technical architecture design in subscription-based CPPSS and provides a foundation for further research and industrial application in service-oriented manufacturing.

Appendix

Appendix 1:

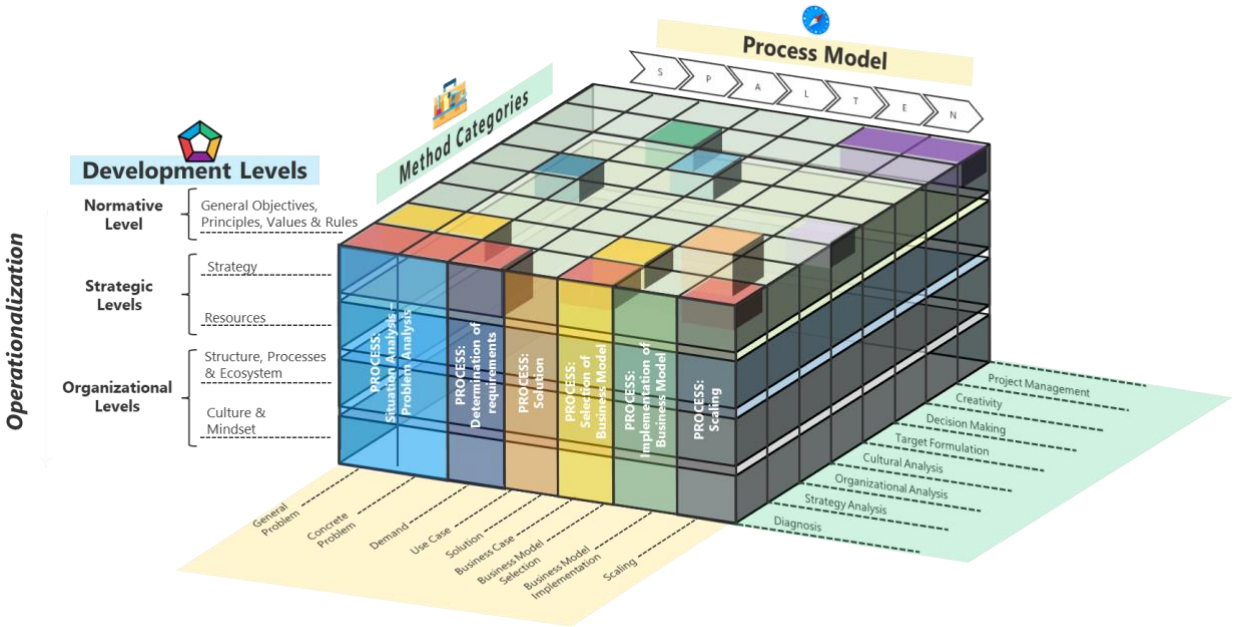


Figure 5: TriProGenix Framework based on [10]

Appendix 2:

M1	Increased customer orientation and interaction	M7	Adapted functional structure and physical architecture
M2	Adapted service offering	M8	Adjusted connectivity and interoperability
M3	Modified monetization logic	M9	Adapted processes
M4	Changed economic performance	M10	Increased process automation and efficiency
M5	Adapted system configuration	M11	Adapted organizational structures
M6	Enhanced quality and safety	M12	Changed external environmental factors

Figure 6: Characteristics of subscription-based CPPSS

Appendix 3:

		Elements of the architecture				
		Actuation	Edge gateway	Data storage	...	**Sum** (Influence)
Consolidated characteristics	Adapted lifecycle Logic	2	0	1	0	3
	Adapted monetization	0	1	1	1	3
	Increased data collection and analysis	0	1	2	2	5
	...	0	0	0	2	2
	Sum (Influenced)	2	2	3	5	

Figure 7: Exemplary mapping of consolidated characteristics against elements of the reference architecture

Appendix 4:

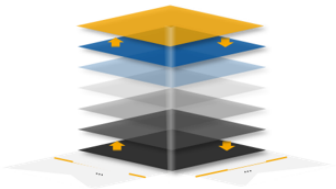
ID 1.1.1 laufende Nummer #1	**Over-The-Air Functionalities**	
Layer	Edge & Connectivity	**Scheme**
Object	Device Control Unit	
Beschreibung		
Description		
Source XYZ		**Difficulty** ■■■■□ **Effort** ■■■■■

Figure 8: Exemplary design brief

References

[1] Wirth, J., Schneider, M., Hanselmann, L., Fink, K., Nebauer, S., Bauernhansl, T., 2024. An Exploratory Analysis of the Current Status and Potential of Service-Oriented and Data-Driven Business Models within the Sheet Metal Working Sector: Insights from Interview-Based Research in Small and Medium-Sized Enterprises. Sustainability 16 (7), 2603.

[2] Wirth, J., Seidel, T., Bauernhansl, T., 2025. Development Of A Framework For A Structural Analyzation And Examination Regarding The Need Of Serviceoriented And Datadriven Business Models In Mechanical- And Plant-Engineering. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[3] McCarthy, D.M., Fader, P.S., Hardie, B.G., 2017. Valuing Subscription-Based Businesses Using Publicly Disclosed Customer Data. Journal of Marketing 81 (1), 17–35.

[4] Schuh, G., Boshof, J., Dölle, C., Kelzenberg, C., Tittel, J., 2020. Subskriptionsmodelle im Maschinen- und Anlagenbau, in: , Internet of Production - Turning Data into Value. Fraunhofer-Gesellschaft, Aachen, pp. 282–301.

[5] Accenture, 2025. What Industrials can learn from other industries regarding Servitization.

[6] IFS, A., 2025. State of Service 2025: Manufacturing Transformation Report.

[7] Ardolino, M., Sala, R., Scalvini, L., Sajjad, S.K., Adrodegari, F., Pezzotta, G., 2024. Barriers and Challenges Toward the Servitization of the Machinery Sector: Evidence from Theory and Practice, in: Thürer, M., Riedel, R., Cieminski, G. von, Romero, D. (Eds.), Advances in Production Management Systems. Production Management Systems for Volatile, Uncertain, Complex, and Ambiguous Environments, vol. 728. Springer Nature Switzerland, Cham, pp. 275–289.

[8] Martinez, V., Bastl, M., Kingston, J., Evans, S., 2010. Challenges in transforming manufacturing organisations into product-service providers. Journal of Manufacturing Technology Management 21 (4), 449–469.

[9] Wirth, J., Höding, M., Hanselmann, L., Schöllhammer, O., Bauernhansl, T., Puchan, J., 2024. Development of a methodology for the derivation of technical requirements for (cyber-) physical product-service systems in service-oriented business models. Procedia CIRP 128, 816–821.

[10] Wirth, J., Erb, S., Hoffmann, F., Metternich, J., Bauernhansl, T., 2024. Framework for the Introduction of Data-driven and Service-oriented Business Models in Mechanical and Plant Engineering Companies. Procedia CIRP 130, 1124–1129.

[11] Wirth, J., Hermann, M., Bauernhansl, T., 2025. Navigating towards New Businesses: An Implementation Framework for Data-Driven and Service-Oriented Business Models. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[12] Aurich, J.C., Clement, M.H. (Eds.), 2010. Produkt-Service Systeme. Springer Berlin Heidelberg, Berlin, Heidelberg.

[13] Goedkoop, M.J., van Halen, C.J., te Riele, H.R., Rommens, P.J., 1999. Product Service systems, Ecological and Economic Basis.

[14] Verein Deutscher Ingenieure e.V., 2021. Entwicklung mechatronischer cyber-physischer Systeme. Beuth Verlag ICS 03.100.40, 31.220.01, 39.020.

[15] Rizvi, M.A.K., Chew, E., 2018. TOWARDS SYSTEMATIC DESIGN OF CYBER-PHYSICAL PRODUCT-SERVICE SYSTEMS, in: Proceedings of the DESIGN 2018 15th International Design Conference. 15th International Design Conference. May, 21-24, 2018. Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia; The Design Society, Glasgow, UK, pp. 2961–2974.

[16] Vandermerwe, S., Rada, J., 1988. Servitization of business: Adding value by adding services. European Management Journal 6 (4), 314–324.

[17] Bruhn, M., Hadwich, K. (Eds.), 2017. Dienstleistungen 4.0: Konzepte, Methoden, Instrumente. Springer Gabler, Wiesbaden, Germany, 1 online resource.

[18] Chen, Y., Visnjic, I., Parida, V., Zhang, Z., 2021. On the road to digital servitization – The (dis)continuous interplay between business model and digital technology. IJOPM 41 (5), 694–722.

[19] Hildenbrand, K., 2006. Strategisches Dienstleistungsmanagement in produzierenden Unternehmen. Dissertationsschrift, St. Gallen.

[20] Lerch, C., Gotsch, M., 2015. Digitalized Product-Service Systems in Manufacturing Firms: A Case Study Analysis. Research-Technology Management 58 (5), 45–52.

[21] Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., Baines, T., 2019. Digital servitization business models in ecosystems: A theory of the firm. Journal of Business Research 104, 380–392.

[22] Schallmo, D.R.A., 2018. Geschäftsmodelle erfolgreich entwickeln und implementieren. Springer Berlin Heidelberg, Berlin, Heidelberg.

[23] Gassmann, O., Frankenberger, K., Csik, M., 2014. The business model navigator : 55 models that will revolutionise your business. Pearson.

[24] Schuh, G., Frank, J., Jussen, P., Rix, C., 2018. Monetizing Industry 4.0: Design Principles for Subscription Business in the Manufacturing Industry, in: IEEE International Conference on Engineering, T.a.I. (Ed.), Conference proceedings ICE/IEEE ITMC. 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC) : Stuttgart, 17.06.-20.06.2018. IEEE, Piscataway, NJ.

[25] Liu, Y., Gützlaff, A., Cremer, S., Grbev, T., Schuh, G., 2021. Design of Tailored Subscription Business Models: A Guide for Machinery and Equipment Manufacturers, in: Behrens, B.-A., Brosius, A., Hintze, W., Ihlenfeldt, S., Wulfsberg, J.P. (Eds.), Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 717–728.

[26] Weber, W., Kabst, R., Baum, M., 2018. Einführung in die Betriebswirtschaftslehre. Springer Fachmedien Wiesbaden, Wiesbaden.

[27] Ebi, M. Implementierung von Subskriptionsmodellen im Maschinen- und Anlagenbau, 284 pp.

[28] Alaluss, M., Drechsler, C., Kurth, R., Mauersberger, A., Ihlenfeldt, S., Marré, M., Labs, R., 2022. Usage-based leasing of complex manufacturing systems: A method to transform current ownership-based into pay-per-use business models. Procedia CIRP 107, 1238–1244.

[29] Mennenga, M., Rogall, C., Yang, C.-J., Wölper, J., Herrmann, C., Thiede, S., 2020. Architecture and development approach for integrated cyber-physical production-service systems (CPPSS). Procedia CIRP 90, 742–747.

[30] Hicking, J., 2020. Spezifikation von intelligenten Produkten im Maschinenbau, 1. Auflage ed. Apprimus Wissenschaftsverlag, Aachen, 412 pp.

[31] Wirth, J., Kallen, P., Bauernhansl, T., 2025. Adaptation and Operationalization of a Process Model for the Introduction of Service-Oriented and Data-Driven Business Models in Mechanical and Plant Engineering. Procedia CIRP 134, 384–389.

[32] Wirth, J., Kallen, P., Bauernhansl, T., 2025. Adaptation and Operationalization of a Process Model for the Introduction of Service-Oriented and Data-Driven Business Models in Mechanical and Plant Engineering. Procedia CIRP 134, 384–389.

[33] Kaplan, R.S., Norton, D.P., 2008. The execution premium: Linking strategy to operations for competitive advantage. Harvard Business Press, Boston Mass., xiii, 320.

[34] Nebauer, S., Schneider, M., Schöllhammer, O., Rauh, L., Schel, D., Bauernhansl, T., Evcenko, D., Kett, H.J., Wirth, J., Bauer, K., Sawilla, I., Höding, M., 2023. Serviceorientierte Produktion der Zukunft.

Biography

Marius Höding (*1992) has been a doctoral candidate at WZL Machine Tool Laboratory (WZL) of RWTH Aachen since 2022 and works as a Technology Strategy & Advisory Manager at Accenture. His research interest is the impact of business model decisions on the design of subscription-based product-service systems. In the research project X-Forge, Mr. Höding was responsible for the development of an overarching process framework for Everything-as-a-Service (EaaS).

Jonas Wirth (*1993) has been a doctoral candidate at Fraunhofer IPA/ IFF, of the University of Stuttgart in cooperation with TRUMPF SE + Co. KG since 2022. His research interest is the scale-up of service-oriented business models in value networks. In the research project Factory-X, Mr. Wirth was responsible for the business model project.

Sarah Ekanayake (*1989) completed her PhD in 2018 at WZL Machine Tool Laboratory (WZL) of RWTH Aachen. After holding leadership positions in industry at Accenture where she worked on the research project X-Forge and Hilti AG, she is a lecturer in Digitalization and Production Management at IPEK of University of Applied Sciences OST in Rapperwil, Switzerland since 2025. Her research focuses on operational excellence, digitalization and global production network.

Thomas Prefi (*1964) has been adjunct professor at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2008 at the Chair of Intelligence in Quality Sensing.

Thomas Bauernhansl (*1969) has been Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since September 2011. He is also Director of the Institute for Industrial Manufacturing and Management IFF at the University of Stuttgart. His scientific focus is on digital and biological transformation, production organisation, factory planning, automation and robotics.

8th Conference on Production Systems and Logistics

Development Of A Process Model For The Selection And Development Of Service-Oriented And Data-Driven Business Models In Mechanical And Plant Engineering

Jonas Wirth[14*], Anka Zoefeld[3], Guenter Bitsch[3], Thomas Bauernhansl[12]

[1] *Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12, 70569 Stuttgart, Germany*
[2] *IFF of University of Stuttgart, Nobelstraße 12, 70569 Stuttgart, Germany*
[3] *ESB Business School Reutlingen, Altenburgerstraße 150, 72762 Reutlingen, Germany*
[4] *TRUMPF, Johann-Maus-Straße 2, 71254 Ditzingen, Germany*

Abstract

A dynamic market environment and technological advancements within the context of Industry 4.0 requires continuous adaptation and realignment of existing strategies and business models for companies in the original equipment manufacturing sector. Particularly, service-oriented and data-driven business models (so. & dd. BM) offer potential competitive advantages. However, their fundamental differences from the previously dominant transactional business models make the introduction of innovative models challenging. Current literature already offers numerous process and categorization models due to the importance of this adaptation. This paper presents a modified process model on so. & dd. BM to fit the need of original equipment manufacturers. Analyzing these categorization models reveals promising business model patterns used as blueprints for the so. & dd. BM development. Integrating these patterns into the development process combines individual service ideas with proven patterns, reducing uncertainty. Therefore, this paper presents a possible process structure for the development of business models with an integrated decision model, which increases the effectiveness of the developed business models, increasing their effectiveness.

Keywords

Servitization; Business Models; Service-oriented; Data-driven; Process Model; TriProGenix

1. Motivation

The German machinery and plant engineering sector is currently facing a significant transformation, which presents new challenges for those Original Equipment Manufacturers (OEM). The crises of recent years have fundamentally altered the market demand, further reducing the already declining market potential in the machinery and plant engineering sector [1]. In response, OEMs are rethinking their traditional business models and are increasingly open to business model innovations [2]. One viable strategy to achieve product differentiation and thus gain a competitive advantage is the implementation of services [3]. Services can enhance both customer acquisition and the creation of long-term customer loyalty [4]. As a result, a paradigm shift often occurs, moving from a product-centered market approach to one supported by services, known as Product-Service Systems (PSS) [5]. This development is commonly referred to as "servitization" [6].

Another driver of the servitization trend is digital technologies, which can be summarized under the term Industry 4.0 [3]. The highly disruptive nature of Industry 4.0 technologies leads to increased competitive pressure, compelling OEMs to rapidly integrate these technologies often times in form of services [7]. The

DOI: https://doi.org/10.15488/20999
ISSN: 2701-6277

introduction of service-oriented and data-driven business models is perceived as a challenge in this context [8]. To enable a consistent, efficient, and future-oriented approach of developing and selecting so. & dd. BM, a process model must be developed to assist companies in navigating this path. Within the scope of this work such a process model and its application will be presented.

2. Fundamentals

2.1 Servitization under the influence of digital technologies

The concept of servitization can be summarized as the combination of products and services. For the purposes of this work, a product is defined as a physical tangible good [9], while a service is considered an activity that does not result in ownership of a tangible item [10]. In recent years, an increasingly dynamic and competitive market environment, particularly for OEMs, has led to a shift from purely product offerings to service-centered solution offerings [11–13].

Today, comprehensive solutions consisting of a bundle of products, services, and knowledge are offered to better convey the overall value of these offerings to customers, thereby increasing customer satisfaction and loyalty [10,9,14,4]. This strategic transformation was first described by Vandermerwe and Rada [4] as servitization. The outcome of this servitization process is referred to as a Product-Service System (PSS) [5]. Given the growing importance of digital technologies, PSS are often enhanced with new technologies and offered as Smart Product-Service Systems (SPSS) [15].

This development is rooted in the scientific discourse known as Digital Servitization, where a PSS is augmented with digital technologies and data collection methods [15,16]. Data and digital technologies provide opportunities to increase value creation and strengthen a company's competitive position through digital and innovative product and service solutions [17,18]. SPSS generate their value through the use of digital product data and can take various forms [6]. The service offerings of smart services range from information provision, condition monitoring, and process control to data-driven predictions, optimizations, and even full automation of processes [19–21].

2.2 Business Model Innovation based on Servitization

The concept of a business model fundamentally describes how a company operates, including the strategies and processes used to create value [6,23,22]. Although there is no universally accepted definition of the term, the concept generally encompasses three dimensions: "value creation," "value delivery," and "value capture" [24,22]. A service-oriented business model in context of application by an OEM can be interpreted as a (S-)PSS with an additional revenue mechanism. The design of new business models and their transformation significantly contributes to long-term business success through differentiation and growth potential, making it of high strategic importance [25,3]. In many cases, the development of new digital technologies is an additional driver for changes in business models [26]. Industry 4.0 technologies, which are based on the processing and use of digital data, are in a continuous development process that forms the basis for modern business models [27,26]. Considering prevailing market conditions and the use of new technologies, SPSS require innovative business models that optimally support realignment. To efficiently develop these business models, a standardized development process is needed that enables companies to quickly develop services and respond flexibly to market demands.

3. Related Research

3.1 Conceptual Basis for the Process Framework

The Business Model Navigator (BMN) by Gassmann et al. (2021) is intended to support companies in the structured implementation of business model innovations [25]. In this context, over 60 business model

patterns (BMP) have been identified, which can serve as templates for redesigning a company's business model [25]. For many companies, these categorized BMP serve as a basis for innovatively advancing their own business models.

The term "business model pattern" is used exclusively in this paper to refer to the 60 patterns identified by Gassmann et al. (2021) to ensure a precise differentiation between the concepts of "business model" and "business model pattern," the terms are defined as follows. A business model is an individual framework that describes how a company creates, delivers, and captures value [22]. In contrast, a BMP is a categorization of business models. The BMN provides a selection of patterns that exist and can be used as a foundation for development. This paper will explore how, based on these patterns, a process can be developed for creating business models specifically for service-oriented and data-driven contexts.

3.2 Structural Basis for the Process Model

Through a structured literature review, nine models were identified that are suitable for supporting the development and selection process of business models. These models range from supporting the decision-making process to the development of an entire business model.

The Performance Improvement Process Model by CHRISTENSEN (2018) [28] describes the process from identifying a need to deciding on a solution and generating ideas. This model is a generic problem-solving model and can therefore be used assessing a customer problem without narrowing down the possible solution.

The process model by DESSINGER et al. (2012) [29] titled "Human Performance Technology Model" addresses the development of the solution in addition to the decision-making process and idea generation. Still also addressing problems in general the range of possible outcomes is not limited to service-oriented and data-driven Business Models.

The BOSCH Innovation Framework (2024) similarly outlines three procedural steps of evaluating a problem to the development of a fitting solution [30]. Though having a more simplified depiction of those process steps.

The Value Proposition Canvas by OSTERWALDER and PIGNEUR (2014) is the first model on this list, which is based on the assumption that the problem has already been evaluated and a solution is to be developed. This model being centred around the value a customer wants to receive, it can be used for the idea generation of business models [31].

Table 1: Analysis of existing models

Existing Models	Process Phases *	formal		content		
		Is it a Process Model?	Simplicity & comprehensibility	Process Triggers	offers various starting points	Integration of BMP
Business Model Canvas *Osterwalder & Pigneur*	4	◑	●	○	◑	○
EaaS Canvas *Fraunhofer IAO*	4	◑	●	○	◔	○
Needs Assessment PIPM (Macro) *Christensen*	1,2	●	◕	○	○	○
BMIP *Adrodegari, Saccani & Pashou*	2,3	●	◑	○	◔	○
Antecendents and Outcomes BMI *Böttcher & Weking*	3	◔	●	●	○	○
HPT (Change Management) *Dessinger, Mosseley & van Tiem*	1,2,3	●	◕	○	○	○
Value Proposition Canvas *Osterwalder & Pigneur*	2	◔	●	◔	●	○
Monetization Canvas *VDMA*	3,4	●	◕	●	◑	○
Bosch Innovation Framework *Bosch Consulting*	1,2,3	●	●	○	○	○

fullfilled: ○ not ◔ hardly ◑ parcially ◕ mostly ● completely

1 BM Needs Assessment 2 Idea generation 3 Elaboration: Idea to BM 4 Depiction of BM

The Business Model Innovation Process by ADRODEGARI et al. (2017) also addresses the idea generation but specifically focuses on the development of a business model [32]. It is kept simplistic with four steps to select and develop a business model with firstly the “idea generation” followed by “future state definition” over “gap analysis” to “action definition” [32].

The Business Model Canvas by OSTERWALDER and PIGNEUR (2011) can depict a business model in its facets [22]. It can be used to communicate the most important components of an already existing business

model. Though the model not depicting the development process or any subsequent processes, it can be used as a tool during the definition of the business model.

The EaaS-Canvas, on the other hand, is an extension of the Business Model Canvas and was developed by KETT et al. (2023) for use cases in the Equipment as a Service environment [33]. “Equipment as a Service” being a SPSS fitting for service-oriented and data-driven business models the specifications of this model were of guidance for the development of a process model for service-oriented and data-driven business models.

The Monetarization Canvas by VDMA (2023) builds on the Business Model Canvas by offering a more process-oriented representation of business models, and it can also be utilized for developing new business models [34].

The presented models were analysed based on formal and content-related requirements, as shown in Table 1. The analysis reveals that a comprehensive development process (process phases 2, 3) for business models does not yet exist. By combining existing models for representing business models with general process models, a process model for the development of so. & dd. BM will be developed.

4. Research Methodology

The Design Science Research (DSR) was chosen as the scientific, structured approach for the development of the model. DSR originated in Information Systems research as a bridge between research findings and research activities [35]. Specifically, DSR is intended to assist in solving complex, application-oriented problems that cannot be addressed by using a single method [36]. Numerous publications and adaptations exist for the DSR approach. This publication draws on the approach by Peffers et al. for Design Science Research Methodology (DSRM) [37]. Peffers et al. identified six activities and four possible entry points within the DSRM [37]. This study is based on a problem-centred approach, whereby the activities are carried out sequentially. This approach is depicted in Figure 1.

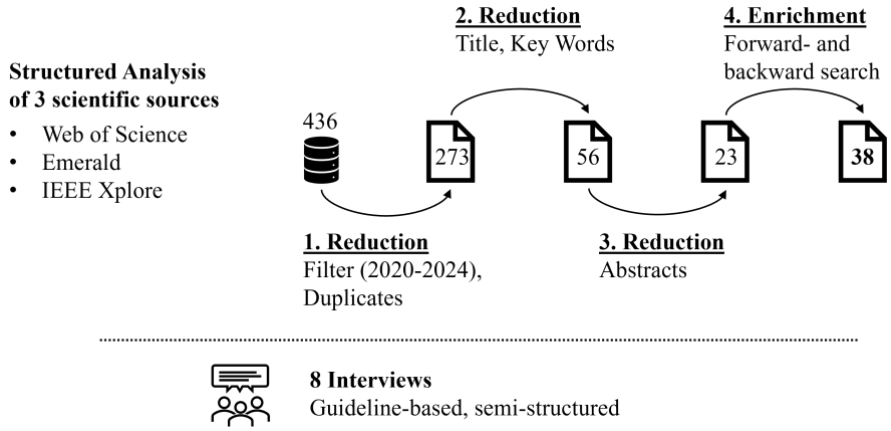


Figure 1: Literature Review Process

Acquiring knowledge in various subfields is essential within these activities. Data collection for this study was divided into three areas, also shown in Figure 1. To establish a solid foundation in the area of service-oriented and data-driven business models, a systematic literature review was conducted using a search string following the methodology of vom Brocke et al. [38]. Building on the theoretical analysis, a qualitative data collection was carried out in the form of eight semi-structured, guideline-based interviews. To methodically underpin the individual steps in developing the decision model, an unstructured literature review was used.

In subsequent research, the process model developed in this paper has been integrated into the three-dimensional TriProGenix (TPG) framework [39] for implementing so. & dd. BM in mechanical and plant

engineering. TPG combines the procedural logic of the model presented here with multi-level development layers (normative, strategic, and organizational) and a method toolbox that provides concrete tools for each process step. In this follow-up work, the original process model and decision logic were adopted as the core navigation path, guiding companies from an initial problem situation toward the design, implementation, and scaling of new service-oriented and data-driven business models.

5. Process Model

Based on existing process models, a process model for business model development in the machinery and plant engineering sector was developed. The development largely relies on the incorporation of practical approaches, which justifies the choice of the BMN as a basis. The model is divided into three steps: a two-part decision-making process and an iterative development process. The structure can be derived from the BMN framework. An overview of the process model presented in this paper can be seen in Figure 2.

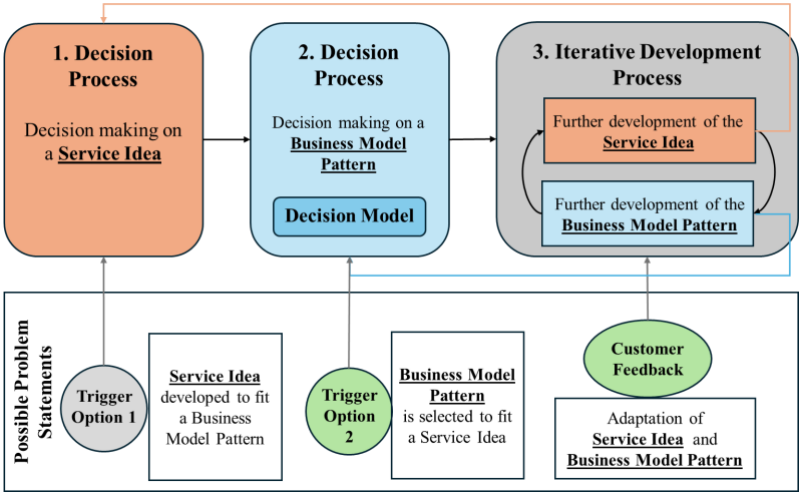


Figure 2: Development Process of a Business Model

5.1 Process Steps

Within the BMN, over 60 BMP have been defined, which companies can use as a reference for developing their own business models (see Section 3.1). The decision to select a particular BMP is often based on an existing service idea. Based on this service idea, which is tailored to the company, a general BMP can be chosen. These two decisions are distinguished by the company-specific decision to adopt a unique service idea and the decision to select a general BMP. In rarer cases, the decision-making process may occur in the reverse order. This sequence arises from the need for a specific BMP, for which a suitable use case is subsequently developed.

Based on the decisions made, the service idea and the BMP are iteratively adjusted to one another. Within the iterative development process, additional development processes, such as the Bosch Innovation Framework, can be utilized. Moreover, the development of service-oriented and data-driven business models requires multiple rounds of customer feedback. These comments also trigger further iterations. The iterative development process concludes with a completed business model, which is thus the outcome of the entire development process. Feedback from the market can also reinitiate the process after the actual business model development process has been completed. This constitutes a business model innovation, wherein the presented process can also be followed.

5.2 Deep dive BMP decision making process

As previously mentioned, the service idea is based on company-specific products and therefore the idea generation can only be supported by tools such as the Value Proposition Canvas but cannot be broadly processualized. In contrast, the BMP remains identical for every company during the decision-making

process. This consistency made it possible to develop a decision model for selecting the appropriate BMP. The different steps included in the decision process are visualized in Figure 3.

The decision model consists of three decision steps:

1. Reduction by Applicability: Firstly, the 60 BMP are filtered based on their applicability to the service idea. All BMP that are not feasible for the use case are excluded from further consideration.

2. Reduction by Requirements: The remaining BMP are evaluated based on their implementation effort. BMP that the company is unable to implement are excluded from consideration.

3. Selection by Impact: In this final evaluation, the impacts of implementing the BMP are considered, and the BMP that offers the best opportunities with the least risks in consideration of the use case is selected.

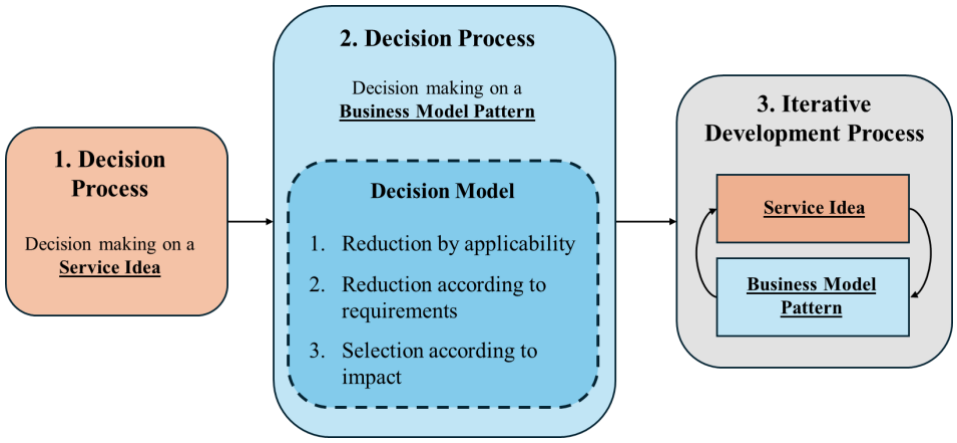


Figure 3: Decision Model of the Decision Process

6. Application of the Framework

6.1 Deep Dive into possible Development Processes

As introduced in the model description, there are various scenarios in which the model can be utilized. The individual possible processes are examined in more detail in this chapter. Depending on the trigger, the process depicted in the process model may proceed differently. Figure 4 illustrates the subprocesses along with their respective triggers.

If the user initiates the process with an existing service idea (a), the next step involves selecting a BMP using the decision-making process. After the decision is made, the BMP is individualized and adapted to the service idea. In the subsequent step, the service is developed from the service idea and tailored to fit the BMP. This mutual adjustment is further developed through multiple iterative cycles and driven by external customer feedback. Once the service and the BMP are comprehensively aligned, the user obtains an individualized, new business model as the process outcome.

In some cases, a company may wish to introduce a specific BMP (b) due to a market trend, for example. The development process then begins with the development or selection of a service applicable to the chosen BMP. After the fundamental development of the service, the BMP must be individualized and adapted to the service. As in the previous subprocess, the service idea and the BMP must be iteratively adjusted through multiple cycles based on customer feedback.

The model is also applicable for adjustments to existing business models (c). Market feedback may reveal the need for a business model innovation. In this case, it must first be decided whether to revise the existing BMP or select a new one. If a new BMP is to be selected for the service, the decision-making process for selecting a BMP can be utilized. Subsequently, the selected BMP must be adapted to the existing service,

and the service must be adjusted to fit the new BMP as well as the changed market demands. Similarly, when revising the BMP, the revised BMP must be adapted to the service, followed by adjusting the service to the BMP. As with the other processes, the adaptation process here is iterative and should be presented to the customer multiple times to incorporate their feedback into the iterative cycles. The result of this process is a business model innovation.

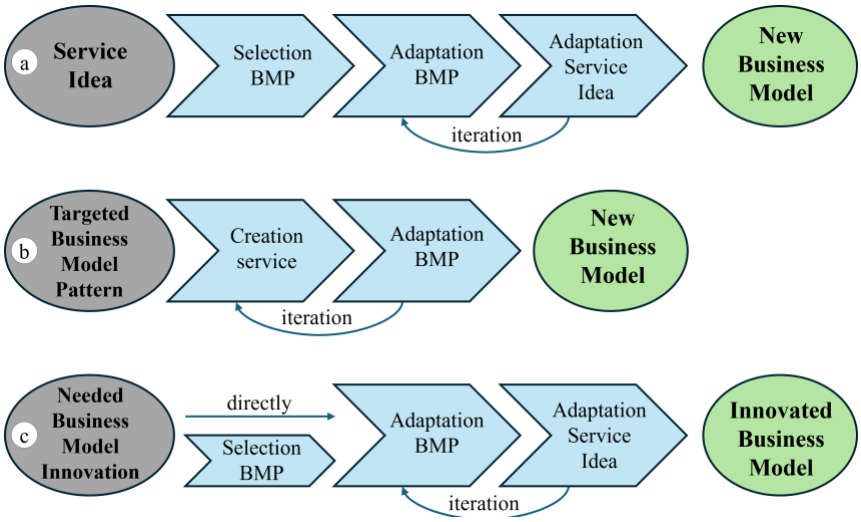


Figure 4: Possible Development Processes

6.2 Example of use

During the validation process, the use case “Shift as a Service” was employed to examine and challenge the proposed process and decision model. The business model “Remote Operation Support”, which underlies the service idea “Shift as a Service” is a real-life business model offered by TRUMPF. The concept of “Shift as a Service” utilizes TRUMPF remote operation services to enable unstaffed production during night shifts. By eliminating the need for on-site machine operators, companies can continue operations during night shifts through remote control conducted by TRUMPF. As a result, automated production can run efficiently without the additional staffing that would otherwise be required [40].

To validate the approach, two workshops were conducted in which experts from TRUMPF were tasked with developing a business model concept for this service idea using the decision model. Because the service idea already existed, the development focused on validating process flow (a), illustrated in Figure 4. The first step of this process flow involves selecting a BMP that aligns with and enhances the existing service idea. By using the integrated decision model, both groups independently selected a hybrid configuration consisting of the BMPs Pay-per-Use and Subscription. In the first step of the decision model, the initial set of 55+ BMPs were reduced to 14 BMP applicable to the “Shift as a Service” use case. As this first decision step relies on objective criteria and factual constraints, divergent outcomes between the two workshop groups were not expected. Consequently, the model was evaluated at this stage primarily with regard to decision speed, comparing decision-making with and without the support of the model. The situation differs in the final two steps of the decision model, in which the requirements and impacts are selected ranked by the participants. In these stages, the usefulness of the model is assessed based on its ability to successfully objectify the evaluation by structuring potentially subjective perspectives. In both workshops, the three BMPs “Pay-Per-Use”, “Performance-Based Contracting” and “Subscription” were initially identified as the most relevant. As previously noted, both workshops ultimately arrived at the same outcome with the support of the decision model. This result indicates that the individual interests and preexisting preferences can be mitigated through the structured evaluation provided by the decision model, thereby fostering more objective decision-making and advancing the potential for innovation. Furthermore, the workshop results closely resemble the actual decision made by the business model development team, highlighting both the effectiveness and time efficiency of the proposed process and decision model. Although the adaption phase

of the process model was not directly validated within the workshops, it reflects the multiple adaptation rounds undertaken by the business model development team in practice. The resemblance between the proposed process model and the real-world development process, as well as their similar outcomes, provide evidence of the model's practical applicability.

The application of the model to "Shift as a Service" demonstrates that the proposed process and decision model functions as a structured navigation system, guiding the transition from a concrete problem to an implementable business model concept. In subsequent work, this logic was generalized and embedded into the TriProGenix framework [39,41], where the process model serves as the procedural backbone across multiple development levels, strategy, resources, structure, and culture and is complemented by dedicated method categories.

7. Conclusion

This paper presented a process model for the general development of a business model based on the BMP of Gassmann et al. (2021). The model is already validated and applicable and can be used as an overview for more detailed planning. By illustrating the triggers, the subprocesses can be presented holistically without distorting the individual processes, thereby providing an initial overview for orientation in further work. The integrated decision-making process functions as additional support during the initial development steps. Therefore, further supporting the goal of this process model to support OEM to innovate their business models.

The findings presented in this paper are built upon various prior research efforts. The process model tailored for business model development integrates approaches from general developmental process models, decision-making models, and insights derived from business model research. The decision to align the process model with the Business Model Navigator facilitates a structured practical application for companies aiming to utilize the business model patterns (BMP) outlined by Gassmann et al. (2021). The objective of the process model is to avoid lengthy and costly development processes through a systematic approach, with the integrated decision-making model significantly contributing to time savings. Thus, this reconfiguration and adaptation of existing theories is value-generating for practical applications.

To further develop this model, it is necessary to assign methodological models to each of the process steps. It may be beneficial to consider the subprocesses separately, as different methods are effective depending on the initial situation. Since the model provides an overview of the entire possible development process, a deeper examination of the subprocesses can also contribute to higher effectiveness.

In a subsequent step, the process model developed in this paper was extended into the TPG framework, which provides a comprehensive navigation system for implementing service-oriented and data-driven business models in mechanical and plant engineering. TPG integrates (i) a six-step procedural model based on a generic problem-solving logic, (ii) multiple development levels ranging from normative and strategic considerations to resources, organizational structures, ecosystems, and cultural aspects, and (iii) a structured method toolbox organized into categories such as diagnosis, strategy analysis, organizational and cultural analysis, target setting, decision-making, creativity, and project management.

Building on the model presented in this article, TPG has been applied and refined in several industrial use cases, including the development of service-oriented business models for Condition Monitoring, the Cutting Assistant, modular automation solutions such as Flex Cell, and the retrospective analysis of EaaS-concepts. These applications demonstrate that the original process and decision model can serve as a robust backbone for comprehensive frameworks that support companies not only in designing but also in implementing and scaling so. & dd. BM

References

[1] Kowalkowski, C., Gebauer, H., Oliva, R., 2017. Service growth in product firms: Past, present, and future. Industrial Marketing Management 60, 82–88.

[2] Dadash Pour, P., Nazzal, M.A., Darras, B.M., 2022. The role of industry 4.0 technologies in overcoming pandemic challenges for the manufacturing sector. Concurrent Engineering 30 (2), 190–205.

[3] Shen, L., Sun, C., Ali, M., 2021. Role of Servitization, Digitalization, and Innovation Performance in Manufacturing Enterprises. Sustainability 13 (9878), 1–17.

[4] Vandermerwe, S., Rada, J., 1988. Servitization of business: Adding value by adding services. European Management Journal 6 (4), 314–324.

[5] Tukker, A., 2004. Eight types of product–service system: eight ways to sustainability? Experiences from SusProNet. Bus Strat Env 13 (4), 246–260.

[6] Koldewey, C., Hemminger, A., Reinhold, J., Gausemeier, J., Dumitrescu, R., Chohan, N., Frank, M., 2022. Aligning strategic position, behavior, and structure for smart service businesses in manufacturing. Technological Forecasting and Social Change (175), 1–10.

[7] Kolagar, M., Parida, V., Sjödin, D., 2022. Ecosystem transformation for digital servitization: A systematic review, integrative framework, and future research agenda. Journal of Business Research 146, 176–200.

[8] Azkan, C., Iggena, L., Gür, I., Möller, F., Otto, B., 2020. A Taxonomy for Data-Driven Services in Manufacturing Industries. Twenty-Fourth Pacific Asia Conference on Information Systems, Dubai, UAE, 2020, 1–14.

[9] Lee, S., Yoo, S., Kim, D., 2016. When is servitization a profitable competitive strategy? International Journal of Production Economics 173, 43–53.

[10] Baines, T.S., Lightfoot, H.W., Benedettini, O., Kay, J.M., Roy, R., Shehab, E., Tiwari, A., 2009. The servitization of manufacturing: A review of literature and reflection on future challenges. Journal of Manufacturing Technology Management 20 (5), 547–567.

[11] Adrodegari, F., Saccani, N., 2017. Business models for the service transformation of industrial firms. The Service Industries Journal 37 (1), 57–83.

[12] West, S., Gaiardelli, P., Rapaccini, M., 2018. Exploring technology-driven service innovation in manufacturing firms through the lens of Service Dominant logic. Ifac-Papersonline 51 (11), 1317–1322.

[13] Wirth, J., Schneider, M., Hanselmann, L., Fink, K., Nebauer, S., Bauernhansl, T., 2024. An Exploratory Analysis of the Current Status and Potential of Service-Oriented and Data-Driven Business Models within the Sheet Metal Working Sector: Insights from Interview-Based Research in Small and Medium-Sized Enterprises. Sustainability 6 (7), 1–19.

[14] Münch, C., Marx, E., Benz, L., Hartmann, E., Matzner, M., 2022. Capabilities of digital servitization: Evidence from the socio-technical systems theory. Technological Forecasting and Social Change 176, 1–16.

[15] Arioli, V., Ruggeri, G., Sala, R., Pirola, F., Pezzotta, G., 2023. A Methodology for the Design and Engineering of Smart Product Service Systems: An Application in the Manufacturing Sector. Sustainability 15 (1), 1–21.

[16] Boucher, X., Pezzotta, G., Pirola, F., Wiesner, S., 2022. Digital technologies to support lifecycle management of smart product-service solutions. Computers in Industry 141, 1–3.

[17] Matusek, M., 2023. Exploitation, Exploration, or Ambidextrousness--An Analysis of the Necessary Conditions for the Success of Digital Servitisation. Sustainability 15 (1), 1–20.

[18] Simonsson, J., Agarwal, G., 2021. Perception of value delivered in digital servitization. Industrial Marketing Management 99, 167–174.

[19] Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., Baines, T., 2019. Digital servitization business models in ecosystems: A theory of the firm. Journal of Business Research 104, 380–392.

[20] Niemann, J., Reich, B., Stöhr, C., 2021. Lean Six Sigma: Methoden zur Produktionsoptimierung. Springer Vieweg, Berlin, Heidelberg.

[21] Rapaccini, M., Adrodegari, F., 2022. Conceptualizing customer value in data-driven services and smart PSS. Computers in Industry, 1–10.

[22] Osterwalder, A., Pigneur, Y., 2011. Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus Verlag.

[23] Marcon, É., Le Dain, M.-A., Frank, A.G., 2022. Designing business models for Industry 4.0 technologies provision: Changes in business dimensions through digital transformation. Technological Forecasting and Social Change 185, 1–15.

[24] Kühne, B., Böhmann, T., 2019. Data-Driven Business Models - Building the Bridge Between Data and Value. Twenty-Seventh European Conference on Information Systems (ECIS), 1–16.

[25] Gassmann, O., Frankenberger, K., Choudury, M. (Eds.), 2021. Geschäftsmodelle entwickeln: 55+ innovative Konzepte mit dem St. Galler Business Model Navigator, 3., überarbeitete und erweiterte Auflage ed. Hanser, München.

[26] Weking, J., Stöcker, M., Kowalkiewicz, M., Böhm, M., Krcmar, H., 2020. Leveraging industry 4.0 – A business model pattern framework. International Journal of Production Economics 225, 1–17.

[27] Opazo-Basáez, M., Vendrell-Herrero, F., Bustinza, O.F., 2022. Digital service innovation: a paradigm shift in technological innovation. JOSM 33 (1), 97–120.

[28] Christensen, B.D., 2018. From Needs Assessment to Needs Analysis. Perf. Improv. 57 (7), 36–44.

[29] Dessinger, J.C., Moseley, J.L., van Tiem, D.M., 2012. Performance improvement/HPT model: Guiding the process. Perf. Improv. 51 (3), 10–17.

[30] Bosch Consulting, 2024. Innovation Framework Design: Let Us Build the Perfect, Tailor-Made Innovation Framework for Your Company.

[31] Osterwalder, A., Pigneur, Y., Bernarda, G., Smith, A., Papadakos, T., 2014. Value proposition design. Wiley, 1-290.

[32] Adrodegari, F., Pashou, T., Saccani, N., 2017. Business Model Innovation: Process and Tools for Service Transformation of Industrial Firms. Procedia CIRP 64, 103–108.

[33] Kett, H.J., Evcenko, D., Falkner, J., 2023. Equipment-as-a-Service (EaaS).

[34] VDMA: Mit Monetarisierungs-Canvas zu mehr Umsatzerfolg., 2023. https://www.vdma.org/viewer/-/v2article/render/75356586. Accessed 14 September 2024.

[35] March, S.T., Smith, G.F., 1995. Design and natural science research on information technology. Decision Support System (15), 251–266.

[36] Nunamarker, J.F., Chen, M., Purdin, T.D.M., 1990. Systems Development in Information Systems Research. Journal of Management Information Systems Winter 1990/1991 (Vol. 7, No. 3), 89–106.

[37] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–77.

[38] vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. 17th European Conference on Information Systems (ECIS), 1–15.

[39] Wirth, J., Erb, S.V., Hoffmann, F., Metternich, J., Bauernhansl, T., 2024. Framework for the Introduction of Data-driven and Service-oriented Business Models in Mechanical and Plant Engineering Companies. 57th CIRP Conference on Manufacturing Systems 2024 (CMS 2024) 2024, 1–6.

[40] Trumpf SE & Co. KG, 2024. Smart Services: Remote Operation Support. Trumpf SE & Co. KG. https://www.trumpf.com/en_US/products/services/services-machines-systems-and-lasers/smart-services/remote-operation-support/. Accessed 23 October 2024.

[41] Wirth, J., Kallen, P., Bauernhansl, T., 2025. Adaptation and Operationalization of a Process Model for the Introduction of Service-Oriented and Data-Driven Business Models in Mechanical and Plant Engineering. Procedia CIRP 134, 384–389.

Biography

Jonas Wirth (*1993) has been a doctoral candidate at Fraunhofer IPA/IFF, of the University of Stuttgart in cooperation with TRUMPF SE + Co. KG since 2022. His research interest is the scale-up of service-oriented business models in value networks. In the research project Factory-X, Mr. Wirth is responsible for the business model project.

Anke Zoefeld (*2001) is a master's student in Industrial Management at Aalen University. During her bachelor's degree in Production Management at Reutlingen University, she collaborated with TRUMPF SE + Co. KG on research into the development of a decision-making model for selecting service-oriented and data-driven business models.

Günter Bitsch is the founder and managing partner of becos GmbH in Stuttgart. He is known as an expert in Manufacturing Execution Systems (MES) and Business Intelligence (BI). In 2010, he earned his doctorate at the University of Siegen. Since 2013, he has taught as an honorary professor at the ESB Business School of Reutlingen University and has been a permanent member of the faculty since the summer semester of 2020.

Thomas Bauernhansl (*1969) has been Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since September 2011. He is also Director of the Institute for Industrial Manufacturing and Management IFF at the University of Stuttgart. His scientific focus is on digital and biological transformation, production organisation, factory planning, automation and robotics.

8th Conference on Production Systems and Logistics

Challenges in Implementing Service-Oriented and Data-Driven Business Models in Mechanical and Plant Engineering: A Literature Review and Empirical Analysis

Jonas Wirth[1,3,4], Tony Scheible[2,3], Yasin Schneider[3,5], Thomas Bauernhansl[1,4]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12, 70569 Stuttgart, Germany*
[2]*University of Stuttgart, 70049 Stuttgart, Germany*
[3]*TRUMPF, Johann-Maus-Straße 2, 71254 Ditzingen, Germany*
[4]*Institute of Industrial Manufacturing and Management (IFF), University of Stuttgart, Germany*
[5]*Hochschule Reutlingen - Fakultät NXT, Reutlingen, Germany*

Abstract

The digital transformation of the mechanical and plant engineering industry poses a significant challenge. As a consequence of increasing market saturation, companies struggle to differentiate themselves from their competitors. In this context, the adoption of service-oriented and data-driven business models may prove an effective approach for companies seeking to maintain their competitive advantage. However, the spread of these innovative business models within the industry has been relatively slow so far. Several factors have already been identified on both the supply and demand sides, but a comprehensive examination of the available data from a theoretical and practical perspective is still missing. This paper consolidates fragmented theoretical and empirical evidence on these barriers by integrating a structured literature review with expert interviews, workshops, and a sector-specific survey in mechanical and plant engineering. The resulting catalogue of challenges and requirements provides a systematic response to the question of why service-oriented and data-driven business models continue to exhibit low adoption rates in practice and identifies the areas where firms must take action. Situated within a broader design-science research program, these findings delineate the empirically grounded problem and requirement space for the subsequent development of the TPG framework, which is designed to support firms in the systematic implementation and scaling of such business models within digital ecosystems.

Keywords

Transformation; Challenges; Service-oriented; Data-driven; Business Models; Digital Ecosystem; TriProGenix

1. Motivation

In recent years, rising costs and intensified global competition have significantly increased the pressure on companies in the machinery and plant engineering sector [1,2]. Under these conditions, differentiation through traditional product offerings alone has become increasingly challenging [3,4]. Service-oriented and data-driven business models (sdBM) promise substantial potential by integrating services with existing products and unlocking new revenue streams [5]. However, despite these advantages, their adoption in practice has been far more limited than anticipated [6,7]. While existing research already addresses Servitization, and data-driven business models, those studies often focus on specific aspects, such as technology enablers, digital infrastructures or isolated organizational barriers. As a result, the existing state of research remains fragmented and a focus on the understanding of the key challenges in Implementing sdBM is still missing. This lack of consolidation and the fundamental shift in business model logic and its

DOI: https://doi.org/10.15488/21000
ISSN: 2701-6277

dimensions, makes it difficult for companies to fully understand the associated challenges and requirements [6]. Against this backdrop, the objective of this paper is to consolidate dispersed theoretical and empirical insights into these challenges and requirements, thereby providing a structured basis for decision-makers in mechanical and plant engineering.

This study is part of a broader design science research program aimed at developing a methodological framework for implementing sdBM within digital ecosystems in mechanical and plant engineering. Within this program, the present paper addresses the first research question: *"What are the key challenges and requirements for mechanical and plant engineering companies when implementing sdBM in digital ecosystems?"* By systematically analyzing and synthesizing evidence from literature, expert interviews, workshops, and a survey, the study delineates the problem and requirement space companies face in practice. These findings provide the foundation for the subsequent development of the TriProGenix (TPG) framework, which is designed to guide companies from recognizing the need for sdBM all the way through implementation and scaling.

2. Fundamentals

The increasing digitalization requires a revaluation or even a complete replacement of existing products, processes, and business models to ensure long-term competitiveness. Addressing these challenges requires a solid understanding of business models and their transformation. Therefore, the following sections will explain and describe the key concepts and changes in business models in detail as well as the phenomenon of Digital Servitization.

2.1 Business Models

Business Models essentially describe how organizations create, deliver, and capture value [8]. They define how a company provides value to its customers and motivates them to pay for it. Furthermore, business models explain how these revenues are transformed into profits [9]. Although there is no single, universally accepted definition of a business model, it is generally understood to encompass three key dimensions: "Value Proposition", "Value Delivery", and "Value Capture" [10].

The "Value Proposition" dimension represents the value a company offers to its customers [10]. It encompasses all processes, resources, and capabilities employed by a company to deliver this value to its customers [11]. To successfully deliver the value proposition, companies must identify key resources, partnerships, activities, and channels. These activities are captured within the "Value Delivery" dimension [10]. The "Value Capture" dimension explains how a business model generates revenue and profits through the delivery of its value proposition, thus achieving economic benefits [12]. Not only are companies' products required to adapt to the challenges in the machinery and plant engineering industry, but their business models must also evolve. While product innovations remain important, they are no longer sufficient on their own to maintain competitiveness in today´s market. Companies increase their potential for success when they innovate their business models. [13]

Business model innovation can thus be seen as a strategic approach to creating a competitive advantage that is difficult for competitors to replace [8]. By innovating their own business models, companies ensure long-term success and growth potential in the market, which is why it holds significant strategic importance [14].

2.2 Digital Servitization and Product Service Systems

Building on the need for business model innovation, an important trend alongside digital transformation is servitization, which Vandermerwe and Rada [13] describe as the development of customer-oriented bundles of products and services incorporated into a company's portfolio. Adding services increases profit margins and strengthens the competitive position, while also enhancing customer satisfaction and loyalty [15,16].

The result of this servitization process is referred to as a Product-Service System (PSS), which includes both a physical product and an intangible service [17–19].

The connection of these two developments, digitalization and servitization, is referred to in the scientific discourse as digital servitization. This provides companies with the opportunity to strengthen their unique value proposition and stand out more clearly in the marketplace [20]. Sjödin et al. [21] define digital servitization as a transformation of a company's processes, capabilities, and offerings to continuously create, deliver, and capture new service value based on digital technologies.

2.3 Service-oriented and Data-driven Business Models

As a logical extension of digital servitization, a transformation towards sdBM is of high relevance for organizations seeking long-term success. Currently, there is no unified definition for sdBM. According to Kühne and Böhmann [22], data-driven business models are characterized by using data as a foundation to create new opportunities for value creation for customers. These business models are closely related to service-oriented business models, which is why data-driven business models can be considered a subcategory of service-oriented business models [23]. Service-oriented business models focus on building long-term customer relationships. In this context, companies develop solutions that are tailored to the individual needs of customers. The value creation logic is increasingly aligned with the customer and their benefits. [24] According to Bruhn and Stauss [25], the aspect of service orientation encompasses aligning a company with the needs of its customers to enhance customer satisfaction and improve customer loyalty.

Taken together, these foundational concepts illustrate the tactical strategic importance for business models to evolve throughout digital transformation. They offer a theoretical basis to understand both the challenges as well as opportunities. This basis is necessary to implement service-oriented and data-driven approaches within mechanical and plant engineering. From an analytical standpoint, these foundations define the dimensions along which the challenges of implementing sdBM can be systematically structured. Specifically, the business-model dimensions of value proposition, value delivery, and value capture, in conjunction with the organizational layers of the firm and its surrounding ecosystem, constitute the lens through which the empirical findings of this paper are interpreted. This perspective is subsequently incorporated into the design of the TPG framework.

3. Methodology

3.1 DSR Methodology

In this study, the Design Science Research (DSR) approach by Hevner [26] is applied. This research methodology focuses on the development and evaluation of artifacts to solve real-world problems in a business context[26]. An artifact can be any form of model, method, or implementation [27]. In the context of this study, the artifact is represented by a table that is developed in several successive steps. Initially, the knowledge base is expanded through a systematic literature review, following the principles outlined by vom Brocke et al. [28], corresponding to the Rigor Cycle of the DSR process. Subsequently, the Relevance Cycle is implemented through the conduction of five workshops and a qualitative secondary analysis of interview data. The collected data are then analyzed and systematically recorded in the table, which constitutes the central DSR artifact.

Within the overarching research program, TPG is conceptualized as a multi-level framework designed to support mechanical and plant engineering firms in the systematic development and implementation of sdBM within digital ecosystems. In this context, the artifact developed in this study, a structured and empirically grounded catalogue of challenges and requirements, constitutes the foundational problem-space knowledge base for subsequent design cycles and informs the TPG framework. The literature review, expert workshops,

qualitative secondary analysis, and survey presented in this paper thus represent the rigor and relevance cycles that address the first research question of the overall project.

4. Data Collection and Analysis

To establish a robust empirical foundation for analyzing the challenges and requirements associated with implementing sdBM, the study adopts a multi-stage data collection and consolidation process (see Figure 1).

1. **Structured literature review**: Establishes the theoretical foundation and identifies initial challenge and requirement statements.
2. **Expert workshops**: Enrich and contextualize these insights with practice-oriented perspectives from diverse actors in mechanical and plant engineering.
3. **Qualitative secondary analysis**: Draws on interview data to further refine and validate practitioner viewpoints.
4. **Data consolidation and coding**: Transfers all identified statements into a unified data table, followed by iterative coding and categorization.
5. **Survey design and validation**: Derives survey items from the consolidated table and validates them through an online survey among industry practitioners.

Figure 1: Data Collection Process

This sequential approach enables **methodological triangulation** and enhances the validity of the resulting catalogue of challenges and requirements. Literature-based findings are systematically contrasted with practitioner perspectives from different organizational roles, while the survey quantifies perceived agreement and relevance. The collected data is iteratively coded, compared, and consolidated into thematic clusters, allowing overlapping findings to be merged and redundancies to be reduced. Challenges and requirements are then assigned to common categories and refined into a structured catalogue. The following subsections provide a detailed description of each step in the data collection and analysis procedure.

4.1 Structured Literature Review

To establish the theoretical foundation for this study, a structured literature review was conducted in accordance with the procedure outlined by vom Brocke et al. [29,30]. The review draws on seven electronic databases, Web of Science, ACM, EconBiz, IEEE Xplore, ScienceDirect, Scopus, and EBSCOhost. Relevant sources were identified using an advanced search string that combines terms related to challenges and requirements, service-oriented and data-driven (or servitization-related) business models, business-model and value-creation concepts, and the manufacturing and mechanical-engineering domain. After duplicate removal, 599 papers were retrieved across the seven databases. Title and abstract screening resulted in the exclusion of 523 papers; the remaining 76 publications underwent full-text assessment, yielding a final sample of 17 relevant articles. During the screening process, predefined criteria were applied, which considerate for instance thematic relevance to service oriented and data driven business models, focus on manufacturing or mechanical engineering, and the presence of explicit statements on challenges or requirements, while purely conceptual or non-domain specific studies were excluded. The review concentrated on extracting statements pertaining to challenges and requirements for implementing sdBM in

manufacturing and mechanical engineering (see Figure 2, search string see appendix A). These statements were subsequently coded and integrated into a consolidated data table, together with findings from the empirical studies described in Section 4.4.

This structured procedure ensures a comprehensive and targeted identification of literature relevant to service-oriented and data-driven business models, as well as their associated challenges and requirements within the mechanical and plant engineering sector.

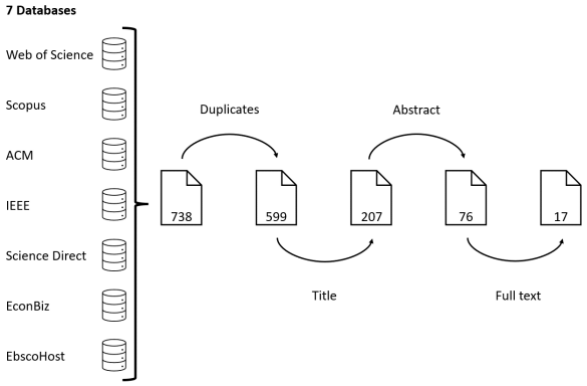


Figure 2: Structured Literature Review (according to vom Brocke [28,30])

4.2 Workshops

To complement the literature-based findings and the extensive interview base analysed in Section 4.3 with a strongly practice-oriented and up-to-date industry perspective, a dedicated workshop format was designed and implemented with experts from the machinery and plant engineering sector. The workshops aimed to identify current challenges and requirements in the context of sdBM and to critically reflect and enrich the insights already derived from interviews.

For the research presented in this paper, workshops were conducted with five industry experts from Wittenstein, DMG Mori, Prenode, Smart Factory Kaiserslautern, and Inovex. Participant selection was guided by the objective of ensuring a heterogeneous and practice-oriented perspective. Accordingly, the group encompassed diverse sectors and roles, including representatives from mechanical-engineering firms, research institutions, software providers, and component manufacturers. The workshops were organized as individual 1:1 sessions, each lasting approximately 45 minutes. A predefined workshop guideline structured the discussions to ensure comparability across sessions. The format was deliberately chosen as an interactive setting that enabled participants to challenge, specify, and complement previously identified challenges and requirements from the perspective of their current projects and organizational context. Statements were documented during and after each session and subsequently integrated into the consolidated data table alongside findings from the literature review and qualitative secondary analysis.

4.3 Qualitative Secondary Analysis

The core empirical backbone of this study is provided by a qualitative secondary analysis of existing interview data. The interviews originate from a broader research program by the author on sdBM in mechanical and plant engineering and were initially conducted for different, yet related, research questions. For the purpose of this paper, those segments of the material addressing the implementation of such business models, perceived barriers and required conditions were systematically re-analysed. This extensive interview base offers a rich, practice-oriented foundation on which the catalogue of challenges and requirements and thus the answer to the research question, is built. The findings from the literature review (Section 4.1) and expert workshops (Section 4.2) complement and sharpen this backbone by adding theoretical perspectives and targeted, up-to-date practitioner insights.

The dataset comprises 82 interview transcripts from 55 interviewees, including both internal and external stakeholders and thus providing a comprehensive view of the topic. Of these, 29 interviewees represent the

focal company, while 26 stem from external organizations such as customers, partners, or other ecosystem actors. Across all interviews, 1,076 codes were identified as relevant for this study because they referred to challenges or requirements associated with service-oriented and data-driven business models. This breadth and depth of coded material enables a differentiated understanding of the obstacles and conditions firms encounter when implementing such business models.

For the qualitative secondary analysis, all interview transcripts were consolidated into a single project. Within this project, existing codes related to service-oriented and data-driven business models, their implementation, and associated challenges or requirements were systematically selected and examined in detail. Where necessary, the original coding was refined and harmonized to ensure consistent categorization across the dataset and comparability with the results from the literature review and workshops.

Relevant passages were paraphrased and transferred, together with contextual information (e.g. stakeholder perspective, company context), into a structured table of interview-based challenge and requirement statements. This interview-based table, combined with the findings from the literature review and workshops, constitutes a key input for the consolidated challenge-and-requirement data table described in Section 4.4, which serves as the central artefact for the subsequent survey and analysis.

4.4 Consolidation into a challenge and requirement data table

Building on the three source-specific tables derived from the structured literature review (Section 4.1), expert workshops (Section 4.2), and qualitative secondary analysis of interviews (Section 4.3), all extracted statements related to challenges and requirements were merged into a consolidated data table. This integrated table constitutes the central artefact of the data collection and analysis phase. For each statement, a concise paraphrased summary was retained to capture its core content consistently and eliminate redundancies in wording across sources.

In the next step, similar statements from different sources were grouped and merged into preliminary sub-categories. These sub-categories were iteratively refined and aggregated into higher-level categories representing major challenge domains, such as organizational and capability-related issues, technological and data-related barriers, customer- and market-oriented challenges, and legal or ecosystem-related aspects. Throughout this process, particular attention was paid to preserving differences in perspective (e.g., internal vs. external stakeholder views) while reducing overlap between categories. The outcome is a structured catalogue of challenge and requirement categories that synthesizes the heterogeneous material into a coherent framework (see Figure 3).

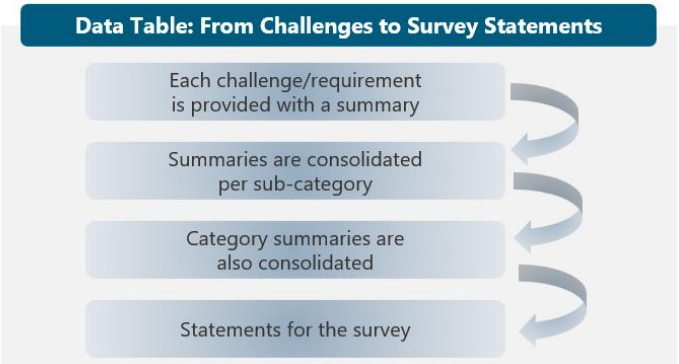


Figure 3: From Challenges to Survey Statements

Based on these categories, concise survey items were formulated to capture the essence of the underlying statements in practice-oriented wording. Ten items were selected to reflect the most frequently mentioned and most relevant challenge areas across all data sources. These items form the basis of the subsequent survey, in which practitioners assess both their level of agreement and the perceived relevance of each challenge. In this way, the consolidated data table not only serves as the core artefact of the qualitative

analysis but also provides the methodological bridge between the rich empirical foundation of interviews, workshops, and literature and the quantitative validation presented in the survey.

Taken together, the multi-stage data collection and analysis procedure establish a robust empirical foundation for this study. The structured literature review, expert workshops, and qualitative secondary analysis of interviews were integrated into a consolidated data table, which was subsequently condensed into major challenge categories and translated into ten survey items. The following chapter builds on this consolidated dataset and the survey results to present and discuss the central challenges and requirements for implementing sdBM in mechanical and plant engineering.

5. Challenges and Requirements for the implementation of sdBM

This chapter presents the central challenges and requirements for implementing sdBM in mechanical and plant engineering. The findings are derived from the consolidated data table described in Section 4.4, which integrates results from the structured literature review, qualitative secondary analysis of interviews, and expert workshops. In total, 1,321 statements were coded and assigned to seven overarching challenge domains: Technical Criteria, Value Creation Network, Capabilities, Organization, Value Proposition, Customer Management, and Framework Conditions (see Figure 4).

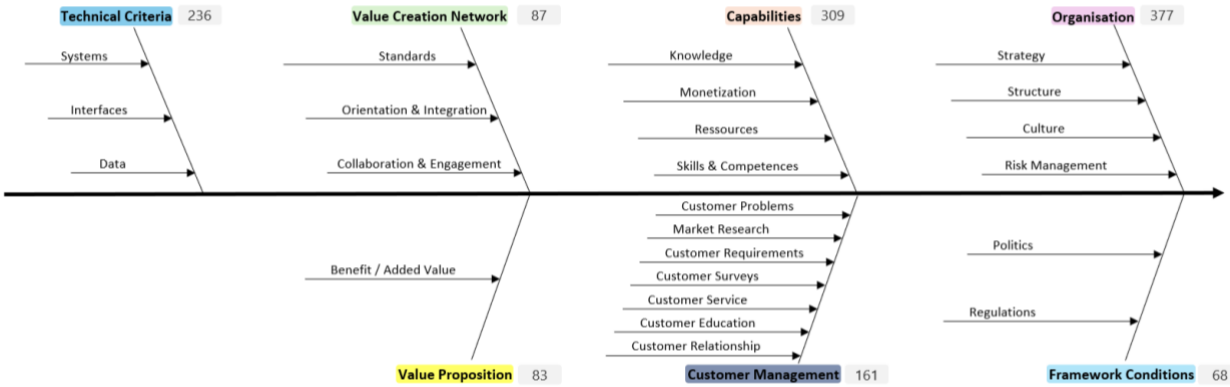


Figure 4: Overview Challenges

5.1 Distribution of Challenges Across Domains

The distribution of statements across these domains highlights the main problem areas. More than half of all statements relate to organizational and capability-related aspects (Organization: 377 statements, 28.5%; Capabilities: 309 statements, 23.4%). Customer Management and Value Proposition account for a further 244 statements (18.5%). In contrast, purely technical criteria such as systems, interfaces, and data infrastructures comprise 236 statements (17.9%), while challenges related to the value-creation network and framework conditions (e.g., standards, regulations) together account for 155 statements (11.7%). Overall, this pattern indicates that sdBM implementation is perceived predominantly as a strategic and organizational challenge with strong customer-facing implications, rather than as a purely technological problem.

5.2 Organizational Capabilities and Mindset

The transition from a product-centric to a service-centric logic requires profound changes in organizational mindset and capabilities. Many firms report a lack of internal expertise in IT, data analytics, and business-model development, limiting their ability to design, implement, and monetize sdBM. This capability gap often results in a strong reliance on external partners and slows down internal learning and implementation processes. At the same time, sdBM often lack clear strategic anchoring within existing structures and steering mechanisms. Issues related to strategy, organizational structure, culture, and risk management feature prominently in the data. Targeted capability building, structured change-management initiatives, continuous training, and clear strategic positioning of sdBM emerge as essential requirements for

supporting this transformation. Overall, the findings indicate that successful sdBM implementation depends less on isolated skill development and more on a coordinated alignment of capabilities, structures, and organizational mindset.

5.3 Collaboration and Governance

The development of sdBM demands close cooperation both within organizational boundaries and across extended value-creation networks. Firms frequently struggle to identify suitable partners, define roles and responsibilities, and establish equitable risk- and value-sharing mechanisms. Standards, orientation and integration across partners, and collaboration and engagement in the network are recurring themes. Legal and regulatory uncertainty, particularly regarding data protection, liability, and political conditions, further amplifies these challenges and increases the perceived risk of engaging in new forms of collaboration. Clear governance structures, transparent communication, and robust contractual and legal frameworks are critical for building resilient and sustainable collaboration settings.

5.4 Technological and Data-Related Barriers

Technological constraints and data-related issues complicate sdBM implementation, summarized in the domain Technical Criteria. Many firms lack the IT infrastructure required to collect, process, and utilize data across heterogeneous machine fleets and systems. Interoperability challenges between equipment from different manufacturers, limited availability and quality of data, and concerns about data security are frequently cited obstacles. Customers often exhibit reluctance to network their machines or share sensitive production and business data. Consequently, reliable technical infrastructures, robust security mechanisms, and clearly defined rules for data usage and ownership constitute indispensable prerequisites for enabling sdBM.

5.5 Customer Acceptance and Value Communication

Customer doubt towards service-oriented and data-driven offerings remains pronounced, particularly when such models involve recurring payments and extensive data sharing. The coded statements highlight issues such as insufficient understanding of customer problems, limited use of systematic market research and customer surveys, and weaknesses in customer service, education, and relationship management. Providers often struggle to articulate the added value of sdBM in a manner that is concrete, comprehensible, and financially compelling. Aligning sales strategies, adapting pricing and incentive systems, and developing clear, measurable value propositions are therefore critical for mitigating customer scepticism and supporting market uptake.

5.6 Synthesis and Implications

Overall, the findings underscore that successful implementation of sdBM in the machinery and plant-engineering sector requires an integrated approach that simultaneously addresses organizational capabilities and mindset, collaboration and governance, technology and data, as well as customer acceptance and value communication. The distribution of coded statements clearly indicates that the principal barriers are perceived at a strategic and organizational level. The majority of obstacles arise from organizational, capability-related, and customer-management factors, while purely technical issues represent only a relatively small fraction of the identified problem space. In other words, sdBM adoption is not primarily constrained by missing technologies but by how firms strategically position these business models, embed them within structures and processes, develop requisite skills, and align organizational and customer interfaces accordingly. Technical infrastructure and data solutions remain important enablers; however, they only become effective when these strategic and organizational implementation issues are addressed systematically. This multi-level, business-model-oriented perspective provides the foundation for the subsequent survey-based validation of the challenge domains and informs the development of a framework that supports firms in systematically navigating the implementation of sdBM.

6. Survey-Based Validation of Theoretical and Practical Insights

6.1 Survey Design

To validate the qualitatively derived challenge domains, an online survey was conducted within the machinery and plant-engineering sector. The ten survey statements were derived from the consolidated data table (Section 4.4) and represent the most frequently mentioned and most relevant challenges and requirements identified through the structured literature review, qualitative secondary analysis of interviews, and expert workshops. The objective of the survey was twofold: (1) to assess whether these statements are perceived as accurate in practice and (2) to determine their perceived relevance. Participants evaluated each statement using a four-point Likert-type scale ("strongly agree", "somewhat agree", "somewhat disagree", "strongly disagree") to avoid neutral responses. Additionally, they rated the relevance of each statement on a five-point scale and completed a ranking task in which all ten statements were ordered by importance. Demographic information (e.g. position within the company, company type and size, age group) was collected to contextualize the results. The survey was distributed internally and externally via e-mail and social media to practitioners in the machinery and plant-engineering sector. Participation was voluntary and anonymous; data were processed in compliance with applicable data-protection regulations. The survey remained open for six weeks and was completed by 121 participants.

6.2 Survey results and discussion

Overall, the survey results confirm the relevance of the challenge domains identified in the qualitative analysis. Across all ten statements, agreement levels are consistently high. For each item, most respondents selected "strongly agree" or "somewhat agree." This indicates that the catalogue of challenges and requirements derived from the literature, interviews, and workshops reflects issues that are broadly recognized in practice. Figure 5 summarizes the survey results by mapping the ten statements according to their average agreement (x-axis) and perceived relevance (y-axis). Challenges related to customer understanding and trust (C2) and customer reluctance to share data (C9) rank highest in both dimensions and are also positioned at the top in the ranking task. Closely related to this, general customer acceptance (C1) is also rated highly in terms of agreement and relevance, which highlights that customers' willingness to embrace sdBM is a fundamental prerequisite for successful implementation. This underscores that customer acceptance, the ability to communicate the value proposition clearly and transparently, and concerns about data security and data sharing are perceived as the most critical barriers to implementing sdBM. In contrast, organizational structure (C5) and legal and regulatory aspects (C10) receive comparatively lower relevance scores and are ranked towards the end of the list, even though agreement with these statements remains positive. Internal processes (C6), employee mindset (C7), and collaboration in value networks (C8) occupy intermediate positions with high agreement but slightly lower relevance ratings than the most customer-related items. Sales-related aspects, such as incentive systems and the complexity of selling sdBM (C3, C4), are also rated as important, reinforcing the observation that commercialization and communication of sdBM pose substantial challenges.

Taken together, the survey results validate the findings from the qualitative data collection. They demonstrate that the main barriers to implementing sdBM are not perceived to lie primarily in technology or regulation, but in customer acceptance, value communication, sales strategies, internal processes, and mindset. Technical and legal conditions are acknowledged as important boundary conditions, yet they are considered less critical than the strategic and organizational implementation issues highlighted in Chapter 5. The survey thus supports the conclusion that sdBM adoption is fundamentally a strategic and organizational transformation challenge with a strong customer-facing component, rather than a purely technological

problem. These insights provide an empirically grounded basis for the development of a framework that enables companies to systematically address these challenges.

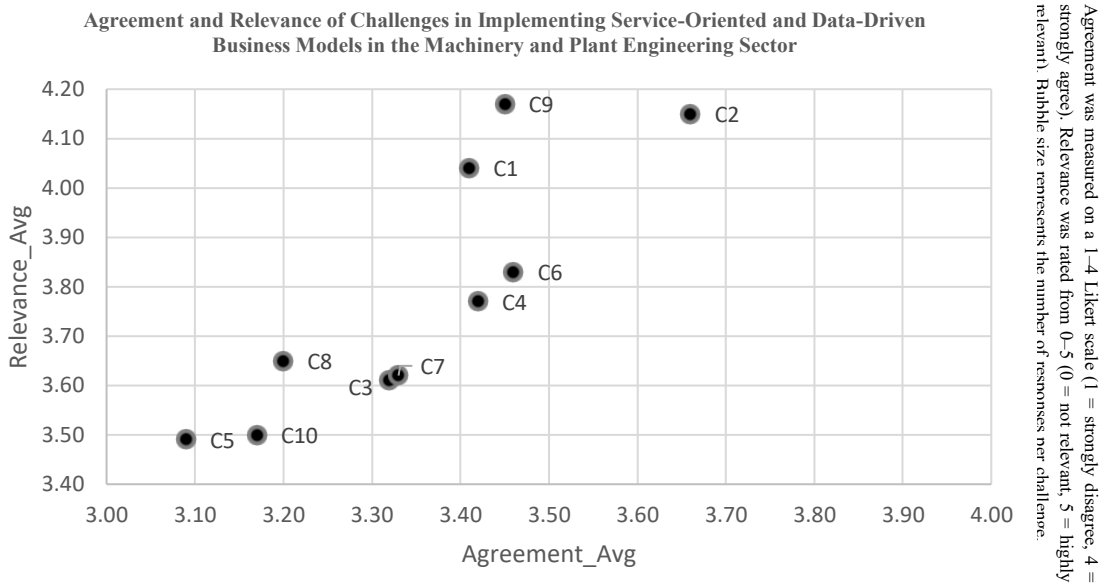


Figure 5: Agreement & Relevance of Challenges in Implementing sdBM in Machinery & Plant Engineering Sector

7. Conclusion and Outlook

This study set out to systematically identify and structure the key challenges and requirements associated with implementing sdBM in mechanical and plant engineering. By combining a structured literature review, qualitative secondary analysis of interviews, expert workshops, and a survey, the research provides a triangulated perspective on the barriers companies encounter when transitioning towards sdBM. The findings reveal that these barriers extend far beyond purely technological issues. While deficiencies in IT infrastructure, data quality, and interoperability remain relevant, most identified challenges pertain to strategic and organizational dimensions as well as customer acceptance.

On the organizational level, companies frequently lack capabilities in IT, data analytics, and business-model development, and typically fail to anchor sdBM within a clear strategic direction. In addition, existing organizational structures, cultural systems, and incentive mechanisms do not align with service-based and data-driven logic, thereby hindering internal adoption.

On the customer level, limited understanding of sdBM, skepticism towards data sharing, the uncertainty to share data, and the inability of providers to clearly communicate the value proposition represents major barriers to customer acceptance. The results of the survey reinforce these findings, by indicating that customer-related, internal process-related, or mindset-oriented barriers are more dominant than the legal or regulatory barriers in the context of market acceptance

Taken together, the results indicate that the implementation of sdBM is less influenced by technology than it is by organizations and customer acceptance. Regarding the design science research program, the results address the first research question by providing an empirically grounded and systematically structured catalogue of challenges and requirements for the implementation of sdBMs in digital ecosystems. This catalogue forms the problem-space basis for the TPG framework, which maps the described problem domains into a graphical, multi-layer model. The study is limited by its focus on the machinery and plant-engineering sector and by the specific sample of interview, workshop, and survey participants. Future research should examine the extent to which these challenges manifest in other industries and contexts and evaluate how design artefacts such as TPG can support firms in systematically overcoming these barriers. Despite these limitations, the study provides a structured, empirically validated view of the challenge landscape of sdBM implementation, offering a robust starting point for both further research and practical transformation initiatives.

Appendix

A. Full search string

"TS=((reason OR cause* OR barrier* OR problem* OR hinder* OR issue* OR challenge* OR difficult* OR require* OR opportun*) AND (service-oriented OR data-driven OR servitization OR "*as a Service" OR *aas) AND ("value delivery*" OR "value creation*" OR "business model*" OR "value proposition*" OR "business strategy" NOT "digital twin*" NOT "machine learning") AND ("manufacturing industry*" OR manufacturing-industry OR "mechanical engineering*" OR "manufacturing sector*" OR *metal* OR machine* OR *plant* OR *production* NOT agri* NOT construction*))"*[1]

[1] This search string is adapted to the specific functionality of each database and applied to titles and abstracts.

B. Overview of the challenges & requirements (without details)

Number	Level 1	Level 2	Level 3	Level 4
377	1. Organization	1.1 Strategy	1.1.1 Employee strategy	
			1.1.2 Sales strategy	
			1.1.3 Data strategy	
			1.1.4 Corporate strategy	
			1.1.5 Competitive analysis	
		1.2 Structure		
		1.3 Culture	1.3.1 Mindset	1.3.1.1 Sales mindset
				1.3.1.2 Customer mindset
				1.3.1.3 Employee mindset
				1.3.1.4 Company mindset
				1.3.1.5 Industry mindset
			1.3.2 Willingness to change / Change management	
			1.3.3 Concerns / Fear	
		1.4 Risk management		
309	2. Capabilities	2.1 Knowledge	2.1.1 Expertise	
			2.1.2 Problem understanding	
			2.1.3 Knowledge management	
			2.1.4 Knowledge transfer	
		2.2 Monetization / financial benefit	2.2.1 Financial benefit	
			2.2.2 Service billing	
			2.2.3 Costs	
		2.3 Resources	2.3.1 Financial resources	
			2.3.2 Human resources	
			2.3.3 Infrastructure resources	
			2.3.4 Resource deployment	
			2.3.5 Time resources	
		2.4 Skills and competencies	2.4.1 Sales competence	
			2.4.2 Data / digital competence	
			2.4.3 Analytical capability	
			2.4.4 Innovation capability	
			2.4.5 Customer competence	
			2.4.6 Operational competence	
			2.4.7 Competence development	
87	3. Value Creation Network	3.1 Standards		
		3.2 Alignment & Integration		
		3.3 Collaboration & Engagement		
236	4. Technical criteria	4.1 Systems		
		4.2 Interfaces		
		4.3 Data	4.3.1 Data security	
			4.3.2 Data quality	
			4.3.3 Data volume	
			4.3.4 Data availability	
			4.3.5 Data processing / data management	
			4.3.6 Data sharing	
			4.3.7 Data foundation	
68	5. Framework conditions	5.1 Politics		
		5.2 Regulations		
161	6. Customer management	6.1 Customer problems / needs		
		6.2 Market research		
		6.3 Customer requirements		
		6.4 Customer survey		
		6.5 Customer service		
		6.6 Customer education / acceptance		
		6.7 Customer relationship		
83	7. Value proposition	7.1 Benefit / value added	7.1.1 Financial benefit / value added	

References

[1] Löwen, U., Hartner, F., 2022. Denken in Wertschöpfungsnetzen und Zusammenarbeit in Ökosystemen am Beispiel von Maschinenlieferanten, in: Meier, K.-J., Pfeffer, M. (Eds.), Produktion und Logistik in der digitalen Transformation. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 251–267.

[2] Wirth, J., Seidel, T., Bauernhansl, T., 2025. Development Of A Framework For A Structural Analyzation And Examination Regarding The Need Of Serviceoriented And Datadriven Business Models In Mechanical- And Plant-Engineering. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[3] Soellner, S., Helm, R., Klee, P., Endres, H., 2024. Industrial service innovation: Exploring the transformation process to digital servitization in industrial goods companies. Industrial Marketing Management (117), 288–303.

[4] Wirth, J., Hermann, M., Bauernhansl, T., 2025. Navigating towards New Businesses: An Implementation Framework for Data-Driven and Service-Oriented Business Models. Proceedings of the Conference on Production Systems and Logistics: CPSL 2025.

[5] Chen, Y., Wu, Z., Yi, W., Wang, B., Yao, J., Pei, Z., Chen, J., 2022. Bibliometric Method for Manufacturing Servitization: A Review and Future Research Directions. Sustainability 14 (14), 1–26.

[6] Wirth, J., Schneider, M., Hanselmann, L., Fink, K., Nebauer, S., Bauernhansl, T., 2024. An Exploratory Analysis of the Current Status and Potential of Service-Oriented and Data-Driven Business Models within the Sheet Metal Working Sector: Insights from Interview-Based Research in Small and Medium-Sized Enterprises. Sustainability 6 (7), 1–19.

[7] Wirth, J., Seidel, T., Bauernhansl, T., 2025. Development Of A Framework For A Structural Analyzation And Examination Regarding The Need Of Serviceoriented And Datadriven Business Models In Mechanical- And Plant-Engineering. Hannover : publish-Ing.

[8] Osterwalder, A., Pigneur, Y., 2011. Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus Verlag.

[9] Teece, D.J., 2010. Business Models, Business Strategy and Innovation. Long Range Planning 43 (2), 172–194.

[10] Aas, T.H., Breunig, K.J., Hellström, M.M., Hydle, K.M., 2020. Service-oriented business models in manufacturing in the digital era: Toward a new taxonomy. International Journal of Innovation Management 24 (08), 1–15.

[11] Favoretto, C., Mendes, Glauco Henrique de Sousa, Filho, M.G., Gouvea de Oliveira, M., Ganga, G.M.D., 2022. Digital transformation of business model in manufacturing companies: challenges and research agenda. JBIM 37 (4), 748–767.

[12] Daeyoup, K., Jaeyoung, K., 2015. Business Model Innovation through Value Delivery Differentiation: Multiple Case Studies. Indian Journal of Science and Technology 8 (21).

[13] Gassmann, O., Frankenberger, K., Choudury, M. (Eds.), 2021. Geschäftsmodelle entwickeln: 55+ innovative Konzepte mit dem St. Galler Business Model Navigator, 3., überarbeitete und erweiterte Auflage ed. Hanser, München.

[14] Shen, L., Sun, C., Ali, M., 2021. Role of Servitization, Digitalization, and Innovation Performance in Manufacturing Enterprises. Sustainability 13 (9878), 1–17.

[15] Freije, I., La Calle, A. de, Ugarte, J.V., 2022. Role of supply chain integration in the product innovation capability of servitized manufacturing companies. Technovation 118, 1–14.

[16] Huang, W., Yang, J., Wei, Z., 2020. How Does Servitization Affect Firm Performance? IEEE Trans. Eng. Manage. 69 (6), 2871–2881.

[17] Frank, A.G., Mendes, G.H., Ayala, N.F., Ghezzi, A., 2019. Servitization and Industry 4.0 convergence in the digital transformation of product firms: A business model innovation perspective. Technological Forecasting and Social Change 141, 1–21.

[18] Tukker, A., 2004. Eight types of product–service system: eight ways to sustainability? Experiences from SusProNet. Bus Strat Env 13 (4), 246–260.

[19] Weking, J., Stöcker, M., Kowalkiewicz, M., Böhm, M., Krcmar, H., 2020. Leveraging industry 4.0 – A business model pattern framework. International Journal of Production Economics 225, 107588.

[20] Pezzotta, G., Arioli, V., Adrodegari, F., Rapaccini, M., Saccani, N., Rakic, S., Marjanovic, U., West, S., Stoll, O., Wiesner, S.A., Bertoni, M., Romero, D., Pirola, F., Sala, R., Gaiardelli, P., 2023. The Digital Servitization of Manufacturing Sector: Evidence from a Worldwide Digital Servitization Survey, in: Alfnes, E., Romsdal, A., Strandhagen, J.O., Cieminski, G. von, Romero, D. (Eds.), Advances in Production Management Systems. Production Management Systems for Responsible Manufacturing, Service, and Logistics Futures, vol. 690. Springer Nature Switzerland, Cham, pp. 165–180.

[21] Sjödin, D., Parida, V., Kohtamäki, M., Wincent, J., 2020. An agile co-creation process for digital servitization: A micro-service innovation approach. Journal of Business Research 112, 478–491.

[22] Kühne, B., Böhmann, T., 2019. Data-Driven Business Models - Building the Bridge Between Data and Value. Twenty-Seventh European Conference on Information Systems (ECIS), 1–16.

[23] Zolnowski, A., Christiansen, T., Gudat, J., 2016. Business Model Transformation Patterns of Data-Driven Innovations. Twenty-Fourth European Conference on Information Systems (ECIS).

[24] Böhmann, T., Warg, M., Weiß, P., 2013. Service-orientierte Geschäftsmodelle. Springer Berlin Heidelberg, Berlin, Heidelberg, 160 pp.

[25] Bruhn, M., Stauss, B. (Eds.), 2010. Serviceorientierung im Unternehmen: Forum Dienstleistungsmanagement, 1. Aufl. ed. Gabler, Wiesbaden, 551 pp.

[26] Hevner, A., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research. MIS Quarterly 28 (1), 75–105.

[27] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–77.

[28] vom Brocke, J., Hevner, A., Maedche, A., 2020. Introduction to Design Science Research, in: vom Brocke, J., Hevner, A., Maedche, A. (Eds.), Design Science Research. Cases. Springer International Publishing, Cham, pp. 1–13.

[29] 2019 IEEE International Conference on Software Architecture Companion (ICSA-C). IEEE.

[30] vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. 17th European Conference on Information Systems (ECIS), 1–15.

Biography

Jonas Wirth (*1993) has been a doctoral candidate at Fraunhofer IPA/ IFF, of the University of Stuttgart in cooperation with TRUMPF SE + Co. KG since 2022. His research interest is the scale-up of service-oriented business models in value networks. In the research project Factory-X, Mr. Wirth is responsible for the business model project.

Tony Scheible (*1999) is a student of Technology Management at the University of Stuttgart and has conducted research in collaboration with TRUMPF SE + Co. KG, focusing on the challenges and requirements of service-oriented and data-driven business models in the machinery and plant engineering industry.

Yasin Schneider (*2002) is a student of Industrial Engineering -Sustainable Production and Business at the University of Reutlingen and has conducted research in collaboration with TRUMPF SE + Co. KG, focusing on the challenges and requirements of service-oriented and data-driven business models in the machinery and plant engineering industry.

Thomas Bauernhansl (*1969) has been Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since September 2011. He is also Director of the Institute for Industrial Manufacturing and Management IFF at the University of Stuttgart. His scientific focus is on digital and biological transformation, production organisation, factory planning, automation and robotics.

8th Conference on Production Systems and Logistics

From Digital Twin Towards Shop Floor: Codeless Programming Of Pick-And-Place Robots Through Extended-Reality Interaction

Andoni Rivera-Pinto, Johan Kildal
Fundación Tekniker, Eibar, Spain

Abstract

Configuring pick-and-place robots for machine-tending tasks typically requires several days of expert programming and cell setup. This paper presents a digital-twin approach that enables non-experts to program real robots through extended-reality (XR) interaction. Within an XR environment users position the robot end effector, issue spoken commands for perception and motion planning, and automatically generate executable behaviour-tree programs deployable on a physical robot. While functional execution is demonstrated, future work will focus on quantitative accuracy validation. In a user study, novice participants completed a full six-object pick-and-place sequence in under four minutes. Compared with integrator-based workflows, the approach reduces deployment time by an order of magnitude and supports agile, human-centred automation in smart manufacturing.

Keywords

Codeless robot programming; Digital twin; Machine tending; Human–robot interaction; Agile manufacturing

1. Introduction

Industrial robots are essential for achieving flexibility and productivity in manufacturing, yet their adoption remains limited in small and medium-sized enterprises (SMEs). A major barrier is the difficulty of configuring robots for frequently changing tasks such as part feeding, handling, and inspection. Conventional programming methods require specialised personnel who typically need to halt production, resulting in both downtime and increased integration costs. These constraints are obstacles to the practical feasibility of deploying robots for high-mix, low-volume production operations.

To address these limitations, recent research has explored intuitive human–robot interaction and programming-by-demonstration. Our earlier work [13] introduced a digital-twin-based environment where users could teach robot motions in virtual reality using hand guidance and spoken commands. The present study extends this concept to machine-tending and pick-and-place tasks, presenting an end-to-end workflow in which users program a virtual robot twin through extended-reality (XR) interaction while an executable behaviour-tree (BT) program is automatically generated. The resulting BT has been functionally deployed on a physical robot, although quantitative accuracy benchmarking between virtual and real setups remains future work. We evaluate whether non-expert users can successfully complete a full pick-and-place programming routine through this workflow in a controlled user study.

DOI: https://doi.org/10.15488/21001
ISSN: 2701-6277

2. Related work

Programming by Demonstration (PbD) has long been studied as a strategy to enable non-experts to teach robots new skills without writing code. Classic PbD surveys describe kinaesthetic teaching, teleoperation, and multimodal demonstration techniques [2], and several collaborative robots now offer commercial kPbD functionalities. While effective for teaching basic motions, these methods typically require the robot to be taken offline [16], heavily impacting productivity.

Extended-reality (XR) interfaces have been explored as an alternative for robot programming, enabling users to specify motions or task constraints on virtual or mixed-reality representations of the robot. Prior work has demonstrated VR-based trajectory authoring, teleoperation, and workspace annotation [5]. Our earlier contributions introduced an XR environment where users could teach approach and positioning routines by manipulating a holographic digital twin and issuing spoken commands [13]. Several studies have used AR or MR interfaces for robot programming, enabling trajectory specification, waypoint annotation, or hand-guided spatial manipulation of holographic end-effectors. Examples include AR-based welding-robot programming [10, 11] and MR manipulator programming via holographic waypoints [12]. Further multimodal and digital-twin–oriented approaches explore natural manipulation of holographic robot twins, interaction through gaze and gestures, and HoloLens-based hand-guidance frameworks [3, 4].

Behaviour Trees (BTs) have become a popular representation for modular and reactive robot tasks [15]. However, few approaches exist that frame their use in production-relevant tasks [1]. Several recent works have explored automatic generation of behaviour trees for robotic task execution. For example, BTGenBot proposes generating BTs from textual task descriptions using lightweight language models [8], validating execution in simulation and on a physical robot. Similarly, CoBT generates behaviour trees from a single demonstration for manipulation tasks [9], enabling non-expert programming through demonstration-based learning. Other approaches, such as NL2BT, focus on deriving hierarchical robot skills from natural-language instructions [18] and structuring them into executable BTs. In contrast, the present work integrates XR-based interaction, digital-twin programming, machine vision, ROS2 orchestration, and automatic behaviour-tree generation into a single end-to-end programming workflow that supports functional transfer from the virtual to the physical robot.

3. Current practices to robotize a pick and place task in industry

Pick-and-place operations such as machine tending, component feeding, and part sorting, are among the most widely automated processes in industrial manufacturing. Despite this prevalence, configuring a robot for a new pick-and-place task remains a complex, multi-stage engineering effort that demands specialized expertise and significant time investment. This complexity particularly affects SMEs, which often face technical, organizational, and financial barriers when adopting advanced automation technologies [6].

The setup of a new pick-and-place application typically begins with a comprehensive analysis of the existing manual process. This includes understanding the geometry, variability, and supply method of the parts to be handled; the required cycle times; the positioning of upstream and downstream machines; and any human–robot interactions. Real production environments introduce variable conditions, such as inconsistent part orientation, mixed part types, or imperfect feeding mechanisms that must be considered from the outset.

Based on this process analysis, engineers develop the cell layout and system integration plan. Using simulation tools, they select the robot model, design the cell footprint, validate reachability and movements, and specify grippers, tools, and safety devices. Integration with external systems is also defined at this stage, including PLCs, production management software, and machine-vision components used for part detection or pose estimation. Vision configuration alone often requires careful selection and placement of cameras, calibration of coordinate frames, and design of detection strategies tailored to specific part geometries.

With the cell design complete, the next step consists of offline programming, where an initial version of the robot program is created directly from the simulated environment. This program encodes the high-level task sequence, including approach motions, pick poses, placement patterns, tool actuation and synchronization with sensors. While useful as a starting point, offline programming rarely requires additional on-site commissioning and fine-tuning due to discrepances between the simulated and physical work cell [17].

Once the physical equipment is installed, engineers complete the setup through on-site commissioning and fine-tuning, which may require production interruptions [17]. This involves calibrating base frames and tool frames, verifying the exact pick positions under real lighting and part tolerances, refining trajectories, and resolving issues that arise from mechanical inaccuracies or environmental variation. The programming languages used (such as FANUC TPP, KUKA KRL, ABB RAPID, or URScript) are vendor-specific and require considerable training. For a typical pick-and-place application, dozens of robot positions must be manually taught and validated. Additional effort is required to calibrate cameras, test detection reliability, and adjust grasp strategies until the system achieves stable cycle times and error-free operation.

After commissioning, workers interact with the system mainly through the teach pendant interface to select production "references" or job configurations. These references may contain long lists of parameters, and similar part types often differ in subtle ways, making correct selection difficult for non-specialists. While operators can start and monitor production, they generally cannot modify or create new robot tasks; any change in part type, gripper, or process specification requires reprogramming by trained personnel.

Overall, the engineering workflow (initial task analysis, cell design, offline programming, installation, calibration, and final validation) may require several days to weeks of commissioning effort before a new pick-and-place application becomes fully operational, depending on system complexity and calibration requirements [17]. After that, adapting to changes in production demands is slow, costly, and difficult. For all these reasons, new approaches are needed for non-expert personnel to program robots more accessibly.

4. System architecture

The system architecture integrates extended-reality interaction, natural-language communication, robotic control, and machine vision into a unified framework that enables users to program and validate robotic tasks without writing code. Figure 1 illustrates the main components and information flows. The architecture follows a distributed design in which XR technologies handle user interaction and spatial task definition, while ROS2 orchestrates robotic execution, perception, and communication with the physical cell.

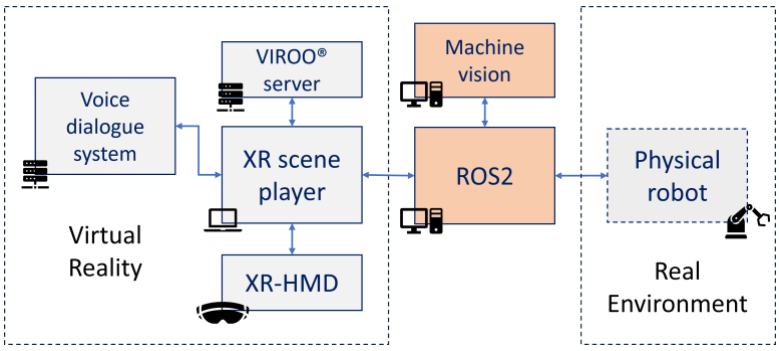


Figure 1: System architecture

At the core of the system is the XR scene player, responsible for executing the immersive virtual environment in which users teach robot behaviours. The XR scene player runs on a workstation connected to a VIROO® (https://virtualwareco.com/viroo/) server, which stores and distributes the XR scenes. The server maintains synchronized state across sessions, enabling the same virtual cell to be accessed concurrently by multiple users when required. The scene player presents the virtual workspace through an XR-HMD, providing

stereoscopic rendering, hand-controller tracking, and spatial manipulation capabilities. Users position virtual markers, define target poses, and issue task commands through direct manipulation within this environment.

Natural-language interaction is enabled through a voice dialogue system, which receives text from transcribed speech and extracts high-level task intents, action targets, and parameter references. The dialogue system operates asynchronously alongside the XR interaction loop and communicates with the XR scene player through lightweight network messages. It interprets user utterances such as “analyse the region,” “plan,” or “place the object,” mapping them to structured actions that the system can execute. This allows declarative control over robot behaviour without requiring exposure to programming syntax or low-level robot commands.

The XR subsystem communicates with ROS2, which functions as the central middleware layer governing both virtual and physical robots. ROS2 manages trajectory planning (via MoveIt 2), state updates, behaviour-tree generation, and robot-agnostic command routing. Actions triggered in the XR scene (such as movement requests, scene queries, or object interactions) are translated into ROS2 service calls or actions. The same ROS2 infrastructure is used for both simulation and real-world deployment, enabling the digital-twin paradigm in which the virtual robot mirrors the kinematic and dynamic characteristics of the physical arm. When interacting with an actual robot, ROS2 forwards joint trajectories, tool commands, and state feedback through manufacturer-specific drivers while preserving a consistent high-level interface.

Perception is provided through a machine-vision subsystem connected to ROS2. This subsystem executes object detection and geometric point extraction relevant in pick-and-place programming, such as part poses. It transmits structured observations to ROS2, which integrates them into motion-planning and execution routines. The integration via ROS2 actions ensures robust handling of asynchronous operations and allows vision data to be consumed both by the XR environment (for visualization) and by the robot controller.

Finally, the architecture connects to a physical robot, which receives planned trajectories and tool commands from ROS2. The physical robot mirrors the virtual behaviour either during execution of a recorded behaviour tree or through direct teleoperation mediated by user actions in the XR scene. Because both virtual and physical robots share the same ROS2 interfaces, switching between simulation and deployment requires no modification of the task logic, provided the virtual scene accurately reflects the real-world configuration.

Overall, the architecture combines immersive XR-based programming, natural-language interaction, robot-agnostic control, and integrated perception into a cohesive system that lowers the complexity of robotic application development.

5. Interaction workflow

The interaction workflow enables users to program a complete pick-and-place task directly within the XR environment by combining spatial manipulation, natural-language commands, and automatic program generation. Inside the virtual cell, the user manipulates a virtual pose marker (a red sphere with an arrow) using the motion-tracked controllers. The user positions this marker above the region of interest, thus defining the desired end-effector pose for viewing the parts scattered on the table.

Once the pose is set, the user triggers motion planning through the spoken command “*plan*”. The planned trajectory is visualized in XR, and, if satisfactory, the user confirms execution with “*execute the plan*”, causing the virtual robot to move its onboard camera to the marked viewpoint. With the robot observing the workspace, the user requests perception by saying “*analyse the region of interest*”. This prompts the machine-vision subsystem to detect all parts and extract their geometric features. The XR environment displays the processed image and detected poses for user verification. The user specifies the desired placement arrangement verbally using natural expressions, such as: *“place the nuts in a matrix of three rows and two columns.”* From this directive, the system automatically works out the placement targets and computes the full pick-and-place sequence, including approach, grasp, transfer, and placement motions

As these interactions occur, the system records the user's actions, pose selections, and spoken commands, and automatically compiles them into a behaviour tree representing the complete robot program. The behaviour tree captures the structure of the task and references the same actions used during interaction. Because both the virtual and physical robots share the same ROS2 interfaces, the resulting program can be executed on the real robot without modification, provided the virtual scene reflects the physical setup. The generated behaviour tree was deployed and executed on the physical robot, successfully completing the full six-object pick-and-place sequence. This demonstrates functional transfer from the digital twin to the shop-floor setup. However, quantitative benchmarking of positional accuracy and calibration deviations between the virtual and physical environments remains future work.

This workflow therefore provides an efficient path from intuitive XR-based task definition to deployable pick-and-place behaviour, greatly reducing the technical barriers associated with traditional robot programming.

6. Evaluation in user study

The system was evaluated through a user study designed to examine whether non-expert participants could program a complete pick-and-place routine using the XR-based workflow and whether the interaction demands remained compatible with rapid deployment in production contexts.

6.1 Participants

Twelve participants were recruited from engineering and research staff and postgraduate students. Five participants were female. All held university degrees, with backgrounds including robotics, computer science, industrial engineering and machine vision. The group showed varied levels of prior experience with XR systems and collaborative robots. However, they were non-experts in industrial robot programming, as none had prior experience with vendor-specific teach-pendant programming or behaviour-tree-based task specification. All participants opted in voluntarily and provided informed consent.

6.2 Procedure

The study took place entirely in the virtual cell recreated in XR (see Figure 2). Participants were first familiarised with the system and then completed one measured trial. The task consisted of programming a collaborative robot to perform a six-object pick-and-place routine: analysing a disordered set of nuts and washers in a region of interest, identifying required parts through a voice-triggered vision analysis, and commanding the robot to place selected nuts into a 3 × 2 matrix arrangement.

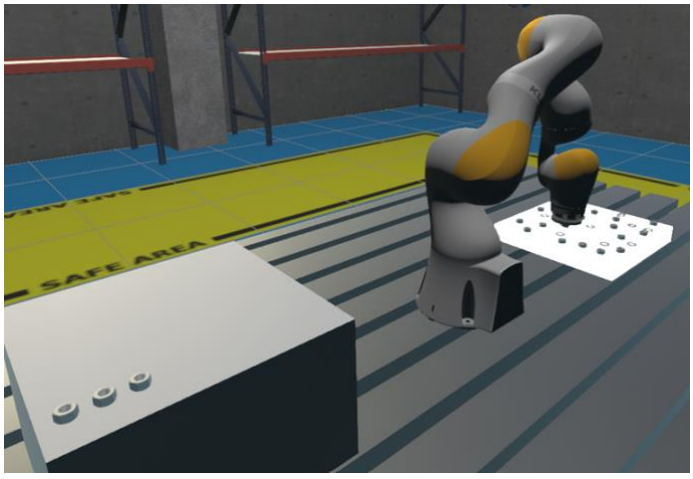


Figure 2: Virtual cell environment. The robot is creating the specified pattern through pick and place.

Participants interacted with the robot through two modalities:

- **Hand-guided positioning** of the end-effector proxy (a holographic sphere with an orientation arrow) above the region of interest (procedure to set the starting position).
- **Spoken instructions** interpreted by the dialogue system, for the execution of actions such as conducting a visual analysis of the unarranged parts, displaying previews of planned actions.

6.3 Quantitative and qualitative metrics

Three quantitative measures were collected:

- **Task completion time**: total time from the start of programming until the virtual execution of the full 2×3 pick-and-place sequence**.**
- **Single Ease Question (SEQ)** [14]**:** 7-point rating of perceived ease of completing the programming task.
- **NASA-TLX (raw)** [7]**:** subjective workload as an index and across six dimensions.

Participants provided comments regarding usability in a semi-structured interview at the end of the study.

6.4 Results

All twelve participants successfully completed the task. The mean programming time was under four minutes (mean 196.3 s, SD 20.9 s). See distribution of observed times in Figure 3. This value includes the entire programming cycle, from initial positioning of the end effector to the robot completing the full six-object placement routine in the virtual twin. An expert user performed the same cycle in approximately 90 seconds, indicating that the interaction overhead decreases with familiarity.

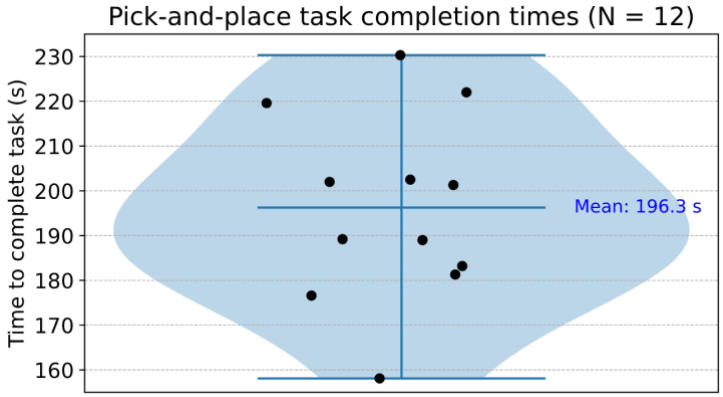


Figure 3: Distribution of completion times for the pick-and-place programming task (N = 12).

Subjective ratings were highly positive. Participants reported high ease of task completion (mean SEQ 5.58/7; see Figure 4). Rated workload was also low (mean NASA-TLX index 3.78/10; see Figure 5).

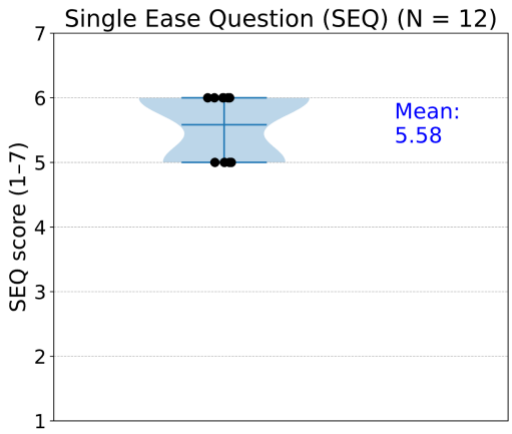


Figure 4. Single Ease Question (SEQ) scores for the pick-and-place programming task (N = 12). Scale ranges from 1 (very difficult) to 7 (very easy)

In the end-of study semi-structured interview, participants consistently reported two sources of difficulty, which contributed to time variance across participants:

- **Grabbing and positioning the holographic end-effector proxy** required more precision than some users initially expected. Minor misalignments occasionally required repositioning.
- **Dialogue-system interpretation errors** occurred when speech was unclear or delivered too quickly, which required repeating spoken instructions.

Across the twelve trials, the dialogue system produced approximately 10 command misinterpretations in total, requiring occasional repetitions of spoken instructions (less than one retry per participant on average). Despite these interaction errors, all participants successfully completed the full six-object pick-and-place routine without external assistance, indicating functional robustness of the overall programming workflow. In the end-of-study interviews, several participants emphasised that natural-language interaction was one of the most intuitive aspects of the workflow.

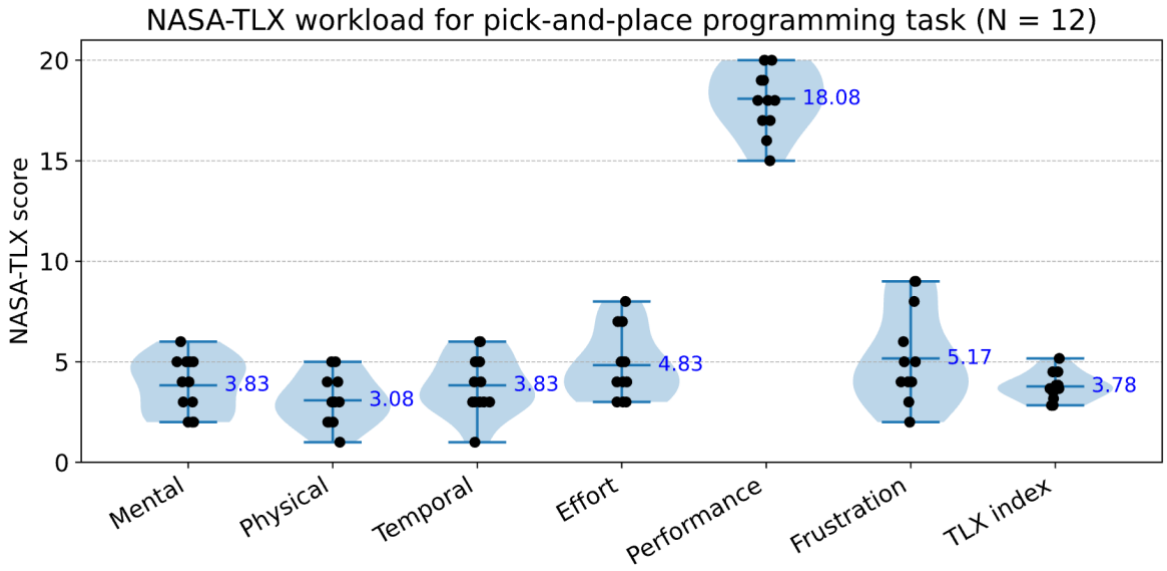


Figure 5: NASA-TLX workload scores for the pick-and-place programming task (N = 12).

7. Discussion

The results show that non-expert users can program a complete pick-and-place routine in a digital-twin environment in under four minutes, a timescale fundamentally different from current industrial practice. As

described earlier, configuring a new pick-and-place application typically requires multiple stages (process analysis, cell design, offline programming, physical commissioning, and final calibration) which often accumulate to more than two weeks before a functional solution is deployed. While the XR workflow presented here does not eliminate the mechanical integration and safety-critical components of this process, it substantially reduces the task-definition and programming effort traditionally performed via simulation software and a teach pendant. The capacity to define robot behaviour interactively and generate an executable behaviour tree aligns well with industrial demands for faster reconfiguration in high-mix environments.

The findings also suggest that the workflow scales well across users with diverse technical backgrounds. Variability in completion times remained relatively small, and subjective workload was low across all NASA-TLX dimensions. This consistency indicates that the interactive system design (combining spatial manipulation, natural-language instructions, and automatic program generation) reduces the cognitive burden associated with traditional robot programming. Users could focus on specifying task intent rather than navigating manufacturer-specific programming languages.

Nevertheless, several limitations must be addressed before broad industrial deployment. Speech-based interaction, although intuitive, showed occasional misinterpretations when commands were delivered rapidly. Improving robustness may be necessary in noisy shop-floor environments. While functional execution of the generated behaviour tree on the physical robot has been demonstrated, the current study does not quantify positional accuracy, calibration sensitivity, or tolerance accumulation between the virtual and the real setups. These aspects will be addressed in future industrial validation studies. It is likely that accurate transfer of virtual task definition to the physical cell will depend on the fidelity of the digital twin, discrepancies in calibration, lighting, or part presentation.

From a deployment perspective, additional considerations include calibration fidelity between virtual and physical coordinate frames, robustness of speech recognition under industrial noise conditions, and compliance with safety standards when integrating with PLC-controlled production lines. Accurate base and tool calibration will be critical to ensure reliable execution, particularly in tasks with tight tolerances. Furthermore, integration into certified industrial environments may require additional safeguards and validation steps beyond those evaluated in this study.

In addition, participants were recruited from engineering and research staff rather than from shop-floor operators. Although they were non-experts in industrial robot programming, further studies with production personnel are required to confirm the generalisability of the approach to typical industrial end-users. Similarly, the current study focused on a single robot and a constrained pick-and-place scenario. Thus, broader validation is also needed across different cell layouts and part types.

Despite these limitations, the approach demonstrates strong potential for reducing reconfiguration time and enabling operators with limited programming expertise to adapt robotic tasks rapidly. This supports ongoing efforts towards agile and human-centred automation in manufacturing.

8. Conclusions and future steps

This work presented a digital-twin based approach to enabling non-experts to program pick-and-place tasks through extended-reality. In the system described, the user utilizes kinaesthetic demonstration of actions in the robotic cell together with providing natural-language instructions for the robot to execute a pick and place task. These interactions with the virtual robot produce executable behaviour-tree programs that have been successfully deployed on a physical robot, while accuracy benchmarking remains future work.

In a controlled user study, participants required under four minutes to create and validate a complete six-object pick and place task programming routine, reporting low workload and high ease of use. These results

indicate that immersive programming interfaces can significantly reduce the effort associated with configuring collaborative robots for frequently changing tasks.

Future work will expand this approach to additional scenarios that are also frequently found in industry. Further research will also explore deployment in live pilot cells to evaluate performance under real production conditions. This will include operator variability, environmental disturbances (e.g., noise, non-ideal lighting conditions) and integration with upstream and downstream production processes. Overall, the results suggest that XR-based programming combined with behaviour-tree execution may provide a promising human-centred and inclusive approach to robotic systems programming.

Acknowledgements

This work has received funding from the European Union under grant agreement no. 101093079 (MASTER).

References

[1] Akkaladevi, S.C., Propst, M., Deshpande, K., Hofmann, M. and Pichler, A. 2024. Towards a behavior tree based robotic skill execution framework for human robot collaboration in industrial assembly. 2024 10th International Conference on Automation, Robotics and Applications (ICARA) (2024), 18–22.

[2] Argall, B.D., Chernova, S., Veloso, M. and Browning, B. 2009. A survey of robot learning from demonstration. Robotics and autonomous systems. 57, 5 (2009), 469–483.

[3] Chu, C.-H. and Lee, I.-Y. 2025. An Experimental Study on the Design of Interaction Modalities in Augmented Reality for Industrial Robot Programming. Journal of Computing and Information Science in Engineering. 25, 5 (2025), 051005.

[4] Chu, C.-H. and Weng, C.-Y. 2024. Experimental analysis of augmented reality interfaces for robot programming by demonstration in manufacturing. Journal of Manufacturing Systems. 74, (2024), 463–476.

[5] Devic, A., Vidakovic, J. and Zivkovic, N. 2024. Development of Standalone Extended-Reality-Supported Interactive Industrial Robot Programming System. Machines. 12, 7 (2024), 480.

[6] Elhusseiny, H.M. and Crispim, J. 2022. SMEs, Barriers and Opportunities on adopting Industry 4.0: A Review. Procedia Computer Science. 196, (2022), 864–871.

[7] Hart, S. and Staveland, L. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Human Mental Workload. P. Hancock and N. Meshkati, eds. Elsevier Science Publishers B.V. 139–183.

[8] Izzo, R.A., Bardaro, G. and Matteucci, M. 2024. Btgenbot: Behavior tree generation for robotic tasks with lightweight llms. 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2024), 9684–9690.

[9] Jain, A., Long, P., Villani, V., Kelleher, J.D. and Leva, M.C. 2024. Cobt: Collaborative programming of behaviour trees from one demonstration for robot manipulation. 2024 IEEE International Conference on Robotics and Automation (ICRA) (2024), 12993–12999.

[10] Ni, D., Nee, A., Ong, S., Li, H., Zhu, C. and Song, A. 2018. Point cloud augmented virtual reality environment with haptic constraints for teleoperation. Transactions of the Institute of Measurement and Control. 40, 15 (Nov. 2018), 4091–4104. https://doi.org/10.1177/0142331217739953.

[11] Ni, D., Yew, A.W.W., Ong, S.K. and Nee, A.Y.C. 2017. Haptic and visual augmented reality interface for programming welding robots. Advances in Manufacturing. 5, 3 (2017), 191–198.

[12] Ostanin, M., Zaitsev, S., Sabirova, A. and Klimchik, A. 2022. Interactive Industrial Robot Programming based on Mixed Reality and Full Hand Tracking. IFAC-PapersOnLine. 55, 10 (2022), 2791–2796.

[13] Rivera-Pinto, A., Kildal, J., Aceta, C., Castillo, A. and Aizpurua, G. 2025. Codeless Teaching of Robot-Arm Positioning in VR. European Robotics Forum 2025 (2025), 42–47.

[14] Sauro, J. 2011. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. CreateSpace Independent Publishing Platform.

[15] Shin, M. and Jung, S. 2024. A Survey of Behavior Tree-Based Task Planning Algorithms for Autonomous Robotic Systems. 2024 15th International Conference on Information and Communication Technology Convergence (ICTC) (2024), 2039–2041.

[16] Sosa-Ceron, A.D., Gonzalez-Hernandez, H.G. and Reyes-Avendaño, J.A. 2022. Learning from Demonstrations in Human–Robot Collaborative Scenarios: A Survey. Robotics. 11, 6 (Dec. 2022), 126. https://doi.org/10.3390/robotics11060126.

[17] Tanz, L. and Daub, R. 2022. Automated commissioning of offline-generated robot programs. Procedia CIRP. 107, (2022), 810–814.

[18] Wang, K., Zhao, Y., Dai, S., Yang, M., He, Y. and Zhang, N. 2024. NL2BT: Learning Hierarchical Robot Skills Represented by Behavior Trees from Natural Language. Cooperative Information Systems. M. Sellami, M.-E. Vidal, B. Van Dongen, W. Gaaloul, and H. Panetto, eds. Springer Nature Switzerland. 366–383.

Biography

Andoni Rivera-Pinto is a researcher in human robot interaction at Tekniker. He obtained his PhD at University of the Basque Country (UPV/EHU) in 2025, with research on multimodal XR based interactions to program robots by demonstration. He is driving technical XR developments in various National and European research projects.

Johan Kildal obtained his PhD in multimodal Human-computer interaction at the University of Glasgow 2008. In the period 2008-2016, he was principal researcher at the Nokia Research Center in Helsinki, developing future interactive solutions and technologies. Since 2016, he is coordinating and contributing to National and European research and development projects, with a focus on XR technologies in the factory of the future.

8th Conference on Production Systems and Logistics

Assessing Data Quality Dimensions And Their Impact On Production Planning and Control

Johannes Jordan[1], Oliver Antons[1], Julia Arlinghaus[2]

[1] *Institute for Engineering of Products and Systems (IEPS), Otto von Guericke University, Magdeburg, Germany*
[2] *Institute for Production and Supply Chain Management, University of St.Gallen, St.Gallen, Switzerland*

Abstract

Accurate, reliable and timely data are the foundation for decision-making in production planning and control (PPC). Relying on an ever-growing number of heterogeneous and interconnected data sources and increasing internal and external uncertainties, software systems are required to gather and process data utilized in decision-making processes. At the same time, data quality is often viewed as given in production research while interconnected production facilities, cyber-physical systems, workers and software systems form a complex network for data exchange and propagation. This paper identifies and categorizes data quality deficiencies relevant to PPC tasks across strategic, tactical and operational planning horizons in order to identify the specific requirements of the tasks. To achieve this, a structured literature review was conducted to study the influence of data quality on the performance of production systems. First, the existing literature was analyzed to determine data quality deficiencies that occur in PPC. Based on the findings the cause-and-effect relationships between data quality and the propagation of these deficiencies through data processing, management, decision-making, and production performance were examined. The results indicate that data quality deficiencies lead to manifold issues, from limiting the effectiveness of PPC tasks to the failure of automated PPC solutions in production after successful prototyping.

Keywords

Production Planning and Control; Data Quality; Manufacturing Performance; Data Acquisition; Data Management

1. Introduction

Manufacturing companies face increasing product variety with shortening product lifecycles. At the same time, there is increasing price competition due to globalization and new competitors in the market. To remain competitive, companies need to ensure their production performance and costs in terms of delivering high-quality products in short and reliable delivery times. To maximize the profitability of a company, they need to minimize process cost, for example by maximizing the utilization of their resources. Furthermore, capital commitment can be reduced through minimized inventories [1]. PPC addresses activities and methods that balance these often-conflicting objectives in manufacturing.

In order to support and improve PPC, in recent years research and industrial practice have focused on collecting more data in real time, as well as on integrating data from multiple sources. This enables the application of data-driven technology for decision-making with an increased degree of automation of PPC tasks. The integration of physical and digital production systems is addressed within research on digital twins and in the broader field of Industry 4.0 [2]. Data quality (DQ) is considered critical, as the quality of the data

DOI: https://doi.org/10.15488/21002
ISSN: 2701-6277

used as input for PPC decisions directly influences their outcomes; missing, incorrect, or uncertain data typically require compensatory buffers such as extended lead times or increased safety stock [3]. With a higher degree of data integration and automation of planning tasks, DQ deficiencies are expected to amplify production performance problems, as errors can affect multiple other planning tasks using the same data source [4]. Nevertheless, the number of publications addressing DQ for PPC is still limited.

An example of cause-and-effect relationships in PPC is the lead time syndrome. The concept is based on the misconception that the gap between planned and actual lead times can be closed by adjusting the planned times to the actual times. The causal relationships can be conceptualized as a cycle. A poor delivery reliability leads planners to increase lead times, which results in earlier order releases. These increase the load on the workstations. The increased load leads to an increase of the queues in front of the workstations, resulting in an extended lead time with increased variability. This decreases the delivery reliability, returning to the starting point of the cycle, causing the syndrome to intensify. With regard to DQ deficiencies, the underlying cause is that the expected lead times are modified by the planner due to performance deficits in the system, causing amplifying problems in the production system [5].

This paper aims to address the question of how specific PPC tasks are affected by DQ and how DQ deficiencies affect production performance. The applied research method is a structured literature review (SLR). The following chapter introduces the theoretical foundation for a research framework. Then the methodology of the SLR will be described in chapter 3. In chapter 4 the results will be presented and discussed before giving a conclusion and outlook in chapter 5.

2. Theoretical Background

2.1 Production Planning and Control

The main PPC activities are the planning and control of resources (capacity), inventory (such as raw material, spare parts, modules and finished products), scheduling and tracking of production activities, responding to unexpected production disturbances and communication with customers and suppliers [6]. PPC can be conceptualized as a closed control loop for achieving production logistics objectives, as illustrated in Figure 1. Target values are determined by the company's objectives and customer demands, which form the basis for planning and controlling all relevant PPC activities. Feedback data collected from the production process enable the identification of deviations between planned and actual values, allowing responses to be initiated in form of adjustments to planning or control activities [7].

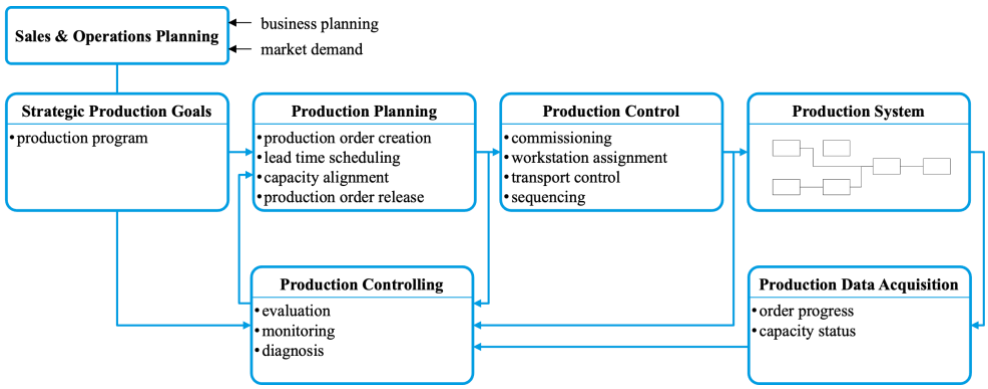


Figure 1: Closed Control Loop of PPC Adapted from [7]

A company's individual PPC process depends on company- and business-environment-specific constraints. Due to the inherent complexity of manufacturing, PPC tasks generally follow a sequential, successive logic that breaks down complex decision-making problems into manageable subtasks [8]. Based on their respective planning horizon, PPC tasks are commonly categorized into strategic production planning tasks

based on forecasted demands, tactical planning tasks based on actual customer orders and operational control tasks to control released production orders [8]. While most subtasks are executed sequentially due to their dependencies, such as strategic planning defining capacity constraints for tactical planning, others may be performed in parallel, for example the scheduling of production orders (tactical production planning) and the sequencing of operations at the workstation level (production control), as both tasks operate within the same planning horizon but rely on different decision scopes. Overall, PPC subtasks differ in scope and information input and output requirements and are closely interconnected [8].

For the definition and conceptualization of PPC tasks, different frameworks exist, using different terminologies for similar tasks. Most common are the Aachen PPC model from Germany [9] and the framework proposed by the American Production and Inventory Control Society (APICS) [6,10]. The *Aachen PPC* framework divides PPC tasks into network tasks, core tasks of production program planning, production requirement planning, procurement and in-house PPC. There are also cross-sectional tasks considering order management and controlling and the cross-functional task of data management. As this paper focuses on single company level, network planning tasks are not relevant. The *APICS framework* categorizes production activities in demand, resource (capacity), planning and execution related tasks.

2.2 Data Quality

The PPC literature distinguishes between transactional data and master data. Transactional data are order-related data or feedback data such as actual inventory levels and start, progress and end dates for production orders as well as updates to estimated completion dates. Master data are parameters used for planning and scheduling production orders. These include, for example, target values for planned times, capacities, stock levels, and batch sizes [1].

In information science a clear distinction between data and information exists. Aamodt defines information as values describing entities like objects, environments or actions and functioning as input for decision-making [11]. Data are the raw material for information as data for itself are observed entities without further meaning or interpretation. Consequently, information is the output of data interpretation [11]. To improve the decision-making in PPC tasks, frameworks for evaluating the quality of information and data exist. The most adopted DQ framework is developed by Wang and Strong, as depicted in Figure 2. Although the authors use the term DQ, their framework is conceptually grounded in information quality, as it evaluates data from the perspective of the data consumer (the entity interpreting the data), the usage context and decision-making relevance. The authors develop a framework that incorporates both intrinsic DQ and dimensions critical to the actual use of the data [12]. Therefore, when referring to DQ this paper focuses both on the quality of data and information. This procedure is supported by relevant literature on information quality also relying on Wang and Strongs framework [12-14].

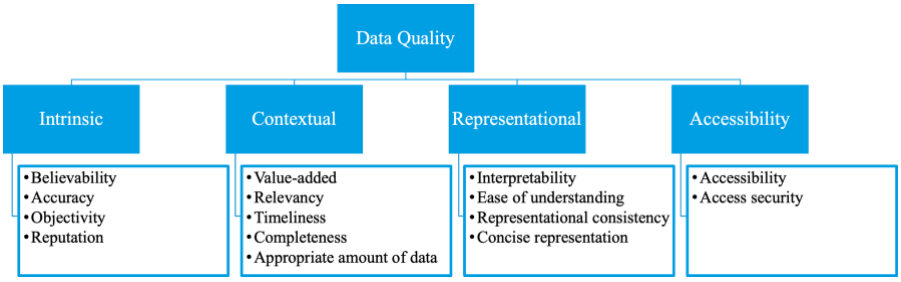


Figure 2: Data Quality Dimensions and Subdimensions [12]

Wang and Strong categorize DQ into the dimensions of intrinsic, contextual, representational and accessibility DQ with subdimensions for each dimension. Intrinsic DQ represents the quality or overall correctness of data itself [12]. Contextual DQ considers the fact that DQ can only be measured according to the requirements of the task that it is relevant for. For example, data entries can be intrinsically correct but not sufficient for the specific task in terms of the timeliness or the amount of data needed. Wang and Strong

define timeliness as the extent to which the age of the data is appropriate for the task being evaluated [12]. Representational DQ is limited due to inconsistent data formats and definitions. Accessibility in terms of DQ represents the possibility of the data consumer being able to access it and the security that people without permission are not able to manipulate or access the data [12].

2.3 Research Gap

To achieve optimal production performance, planners must efficiently utilize control measures, buffers and information [3]. This study aims to examine how DQ affects decision-making in the context of PPC, and, in turn, the achievement of production targets. Therefore, this study aims to contribute to the research questions: *What impact do different data quality deficiencies have on PPC? How do these data quality deficiencies impact production performance?*

3. Methodology

For the study, a SLR was conducted according to Sauer and Seuring [15]. This method is suitable for gaining an overview of the research field and identify further research opportunities. For the design of the search string, the main search terms related to PPC from both the Aachen and APICS frameworks were used. As DQ can only be measured according to the task to be performed, cross-sectional tasks (order, resource/capacity and inventory management) were differentiated based on the time and planning horizon. Using the wording of the APICS framework leads to an adjusted PPC categorization for DQ (left hand side of Table 1) as the basis for the evaluation of existing DQ deficiencies for different PPC tasks. In order to capture the current requirements for DQ in PPC, only publications from 2015 onwards are evaluated. The relevant literature was retrieved from the SCOPUS database. The search resulted in 157 papers as an outcome. The full search string and the process for selecting publications is depicted in Figure 3.

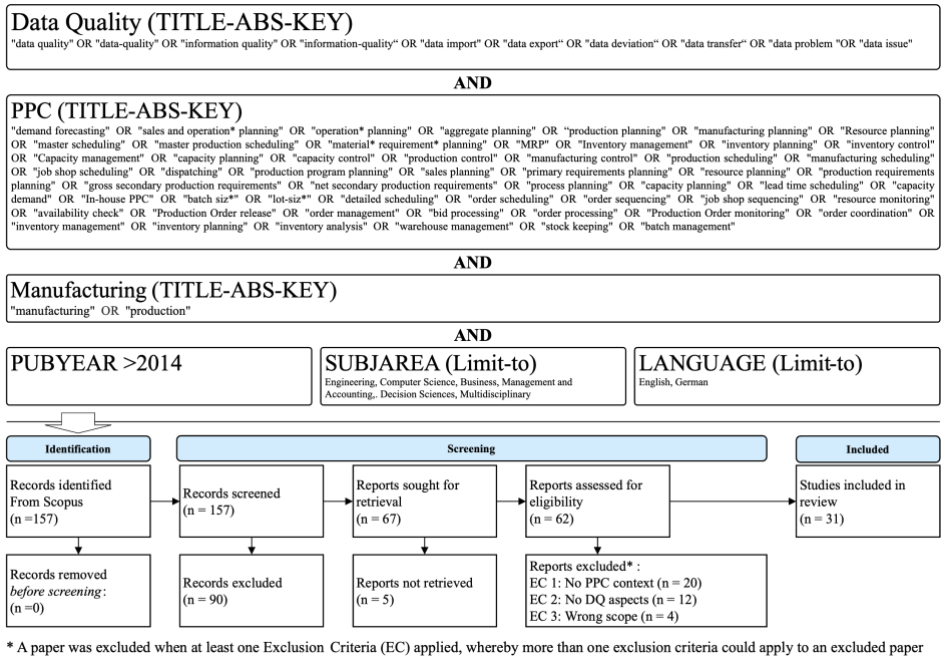


Figure 3: Overview of the Search String and Publication Selection Process

The exclusion criteria can be clustered in not having a PPC or DQ focus and an overall different scope. Mainly publications were excluded due to a missing focus on DQ or PPC. The scope of this study is discrete manufacturing systems at the company level, with a research focus of systems instead of single machines and excluding cross-company data exchange. The abstracts of all papers were screened. This led to the exclusion of 90 papers. Of the 67 remaining papers 62 papers could be retrieved. The full-text reading of the 62 papers led to the exclusion of 31 additional papers and left 31 papers for analysis.

4. Results & Discussion

While there are a multitude of studies on the optimization of inventory, scheduling and other PPC tasks, literature on DQ requirements and problems remain limited. Section 4.1 addresses the first research question by examining how different DQ sub(dimensions) affect PPC subtasks and their respective requirements across different planning horizons. Section 4.2 then analyzes how these DQ deficits impact PPC tasks and overall production performance, thereby addressing the second research question.

4.1 Dimensions of Data Quality Deficiencies in PPC

A review of the extant literature indicates that the planning tasks examined are often described only superficially. Moreover, the field of PPC research lacks a universally recognized framework, as evidenced by the presence of multiple names for similar tasks. These factors make it particularly difficult to categorize tactical production planning tasks. Therefore, in this paper tactical planning activities are categorized by master production scheduling (MPS) and material requirements planning (MRP). Production scheduling and inventory planning are treated as separate categories when MPS and MRP cannot be clearly distinguished. "Non-specific tactical production PP" includes all tactical planning activities that do not fit into any of those categories. For exploratory examination of DQ requirements for different PPC (sub-)tasks, a quantitative frequency analysis for mentions of DQ deficiencies was conducted. These deficiencies were coded according to affected PPC (sub-)tasks and associated DQ (sub-)dimensions. The result of the coding in Table 1 reveals that the accuracy, completeness and timeliness of the data emerge as the most mentioned subdimensions.

Table 1: Evaluation of Data Quality Dimensions That Affect PPC (Sub)Tasks. The PPC (Sub)Tasks Are Depicted With Their Respective Strategic (S), Tactical (T) or Operational (O) Planning Horizons and the Number of Literature References to Data Quality Deficiencies (PPC (Sub)Tasks Without Any Mentions Are Intentionally Excluded)

PPC Tasks	**PPC Subtasks**	**PPC Horizon**			**Data Quality Dimensions**			
		S	T	O	Intrinsic	Contextual	Representational	Accessibility
Production Planning (PP)	non-specific PP				4	5	-	-
	demand forecasting	•			9	1	2	-
	Sales& Operations planning	•			3	-	-	1
	non-specific tactical PP		•		8	6	2	-
	MPS		•		4	8	6	-
	MRP		•		4	11	5	-
	Scheduling		•		9	11	4	2
	inventory planning		•		3	-	-	-
Production Control (PC)	non-specific PC			•	10	16	5	-
	inventory control			•	-	1	-	-

For *operational* tasks, which process large quantities of accurate and up-to-date data, both contextual and intrinsic DQ dimensions are mentioned most. Given the shorter decision-making horizon in operational tasks, the requirements for data timeliness are correspondingly higher. The increased requirement for data timeliness can be addressed through real-time data acquisition.

At *tactical* levels the importance of master DQ in terms of representational DQ subdimensions increases. Also, the accuracy of master data such as the accuracy of planned lead times could be increased by using the feedback data from the shopfloor instead of or in combination with expert knowledge [16]. The integration of shopfloor data can provide valuable information to tactical decision-making. Therefore, the integration of data from different software systems and ensuring efficient information flows between operational workers and planners are possibilities to address these deficiencies.

On a *strategic* level, the overall accuracy (intrinsic DQ) remains the most critical DQ dimension, as the intrinsic correctness of data is a prerequisite. The requirements for timeliness decrease because of the long-term planning horizon. DQ at the strategic planning level could be increased through the integration of

transactional data to enhance the data basis for planning assumptions regarding, for example, capacity demands and production cost calculations.

In the context of smart production, the need for high-quality, high-volume and real-time data is present, especially at the operational level. On the other hand, there is a noticeable trend that the increasing volume of data makes it difficult for decision-makers to identify key influencing factors.

4.2 Propagation of Data Quality Deficiencies in PPC

The identified DQ deficiencies and their consequences demonstrate that such deficiencies occur throughout the data management process from data acquisition and processing to data management and decision-making [17]. To visualize these findings, DQ deficiencies are analyzed with respect to each PPC task and its effect on production performance as a closed control loop (Figure 4). Based on the identified patterns, the following propositions are derived from the SLR, describing the propagation of DQ deficiencies in the control loop.

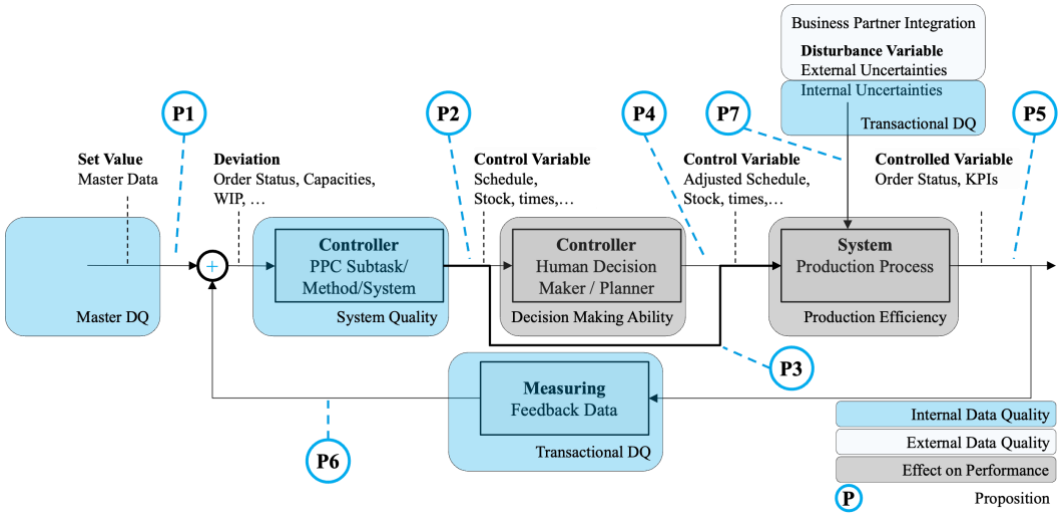


Figure 4: Propagation of Data Quality Deficiencies Within the PPC Control Loop

Proposition 1: *Faulty master data lead to a structural misrepresentation of the production system limiting the quality of the input for the PPC task.*

Master data form the planning basis in PPC, representing the structural parameters of the production system. Keeping master data, for example bills of material [18], correct and up-to-date requires an increasing amount of manual labor as product variants and production complexity increase [19]. The outputs of the PPC tasks are decisions such as batch sizes, order releases or sequencing. This indicates that insufficient input data, in form of intrinsic, contextual and accessibility DQ dimensions of master data lead to suboptimal PPC task outcomes.

Proposition 2: *Correct and timely acquired data can be compromised by a lack of system quality in terms of the processing and management of the data.*

When PPC tasks are conceptualized as a controller, the executing PPC system itself may also be subject to DQ insufficiencies even when the input data are of a high quality. As different software systems (ERP, MES, APS) use data of a different granularity as an input, DQ deficiencies can arise from different calculations, data aggregations or rounding of values and therefore different outputs for similar PPC tasks [17,20,21]. Another challenge is the integration of data sources when collected data is not accessible, for example in distributed data silos [22]. The system quality refers to the representational DQ dimension.

Proposition 3: *Data quality determines the success of fully automated and AI-enabled PPC solutions.*

Low DQ negatively affects the performance of automated PPC tasks [17,23,24,25]. Implementation projects of digital solutions in PPC are highly dependent on the underlying DQ [4,18,26–29,30] and can be delayed

or terminated as a consequence [31]. Rework activities to improve the DQ in production often result in increased manual efforts which set back the automation [17]. Because of the growing amount of collected data higher degrees of automation are becoming increasingly important [32]. Machine learning algorithms need high-quality data as an input. As well-structured and labeled data are scarce, implemented methods are limited in terms of prediction and detection performance [4]. This results in significant manual effort and reduced performance of automated PPC-related algorithms [4]. As many digitization projects such as the implementation of digital twins result in high costs for the acquisition and integration of data, an optimum in terms of cost and performance can only be reached, when the required DQ can be known or measured. This perspective poses a possible future research direction.

Proposition 4: *Humans involved in the PPC process can act both as a stabilizing factor for adjusting PPC decisions to low-quality data and as an additional source of uncertainty. Insufficient DQ induces uncertainty and reduces trust in the data and therefore negatively impact the decision-making ability of human planners.*

Depending on the PPC subtasks degree of automation a second controller is in the control loop in form of a human decision-maker. Inadequate planning results from PPC systems are often compensated by planners through buffers such as increased inventory levels or lead times [33]. While stabilizing operations, these measures result, for instance, in longer delivery times and increased capital commitment [34]. A study describes that with a higher automation decision-makers acceptance towards automatically acquired data decreases as they are not as involved in the data acquisition anymore [17,34]. Known DQ deficiencies and untrustworthy data induce uncertainty and confusion in manual PPC decisions [32]. In context of lead time syndrome, humans transition from passive data consumers to actively influencing its quality by adjusting planning data due to poor adherence to set deadlines. In this case, the planner shifts from a stabilizing factor to an additional source of uncertainty by making subjective risk assessments based on a database perceived as unreliable. If the planned times are not further questioned, this leads to a bias in the data.

Proposition 5: *In multi-layer PPC systems, the negative performance effects of low data quality are reinforced, thereby reducing the stability of the overall production system.*

Considering each PPC subtask as a control loop in overall business operations, decisions made on strategic planning horizons based on low quality input data induce uncertainty and instability in downstream PPC subtasks. DQ deficiencies can also amplify upstream, for example due to a lack of data integration, when strategic planning decisions cannot be made on high-quality actual performance data coming from the shopfloor [35]. On tactical levels uncertainties and low DQ lead to suboptimal decision-making of systems or employees conducting MRP and MPS subtasks [36]. Workers using the data for decision-making are confronted with additional time-consuming efforts to validate the data and acquire or update missing or incorrect data [4,17]. Because of manual and time-consuming tasks, data can be outdated before even being ready for analysis [37].

Proposition 6: *A PPC process is likely to yield suboptimal performance when the current state of the production system is misrepresented in the PPC process due to low data quality.*

A case study of a small and medium-sized metals processing company revealed that of the acquired order start and end dates only 28% of the entries were realistic values without considering the overall correctness of these values [33]. Especially feedback data collected manually from operators lack the DQ needed [16]. Low-quality transactional data interferes with the decision-making process as a disturbance variable as it leads to the real state of the production system being incorrectly represented by the transactional data with regard to intrinsic, contextual or accessibility DQ dimensions. In the worst case, this results in the correct decision being made for the virtual representation of the production systems state. However, this decision is then incorrect for the actual state of the production system deviating from the virtual representation.

Proposition 7: *External and internal uncertainties act as disturbance variables in the PPC control loop, increasing planning variability and reducing production performance.*

As this study focuses on the scope of a single company, DQ deficiencies related to cross-company data integration with suppliers and customers are not considered and uncertainties are not investigated further. These aspects primarily concern the definition and implementation of standardized data formats and interfaces for data exchange. As external and internal uncertainties increase a focus on the internal DQ is expected to increase the robustness towards uncertainty in production.

DQ affects all production performance aspects in terms of quality, cost and time. In terms of quality, the impact of deficits is mainly the quality of the decision by the system or the human planner, referred to in the literature as planning quality/performance [24,35] or robust planning [38]. In terms of inventory, planners may react to DQ deficits by increasing inventories. This can lead to an increased work in progress [35,39–41], decreased capacity availability in terms of machine utilization [33,39] or the deployment of planners and workers involved in increased manual data acquisition and processing [33,34,37]. These factors decrease the production efficiency [39] and ultimately decrease the overall business profitability and the service level [19,32,33,35]. The buffers in form of inventory, buffer times and capacities lead to financial losses that remain hidden and cannot be optimized unless DQ deficiencies can be identified and measured.

4.3 Limitations of the Study

Since the study is based solely on an SLR, the identified DQ insufficiencies (research question one) and propagation mechanisms (research question two) represent propositions derived from studies with different objectives and levels of detail. The quantitative frequency analysis in Section 4.1 does not aim to provide statistical generalization but serves to identify patterns of DQ deficiencies across PPC subtasks. Consequently, the findings require further validation through empirical analyses. The scope limits the study to intracompany DQ aspects without investigation of cross-company data exchange and the influence of uncertainties beyond a company's control. As the classification of DQ deficiencies across PPC tasks and DQ dimensions (Table 1) is based on their frequency of occurrence in the literature, the results reflect general tendencies in existing research and provide guidance for future research directions. The findings do not enable a quantitative assessment of individual DQ (sub-)dimensions on production performance.

5. Contribution and Outlook

This paper provides an overview of the impact of different DQ deficiencies on PPC and their propagation within PPC. Based on an SLR utilizing a quantitative frequency analysis, this study provides a task-specific analysis of DQ deficiencies in PPC by systematically mapping DQ dimensions to strategic, tactical, and operational PPC subtasks. Based on the SLR, accuracy, completeness, and timeliness are identified as the most critical DQ dimensions, with distinct requirements across planning horizons. Beyond the task-level perspective, the paper conceptualizes the propagation of DQ deficiencies within the PPC closed control loop and their increasing impact in highly digitized PPC environments.

As external disturbances such as market uncertainties and competitive pressure in terms of delivery times and costs continue to increase, internal deficiencies in the data acquisition, processing and management become increasingly visible. Possible research directions should further analyze the cause-effect-relationships between DQ deficiencies and production performance to make the impact of DQ measurable. This can contribute to the research aiming to automate data analysis and decision-making to increase the feasibility and profitability of digitization strategies. To analyze occurring DQ deficiencies in industry and successfully transfer AI-enabled PPC methods developed in academic research in industry an exchange of data and methods can contribute to closing the gap between industry and research. Future research should also continue to investigate how DQ deficiencies, such as human-induced DQ deficiencies, can be effectively incorporated into the decision-making process, taking into account increasing uncertainty, data collection, and PPC automation.

Acknowledgements

This research was supported by the state of Saxony-Anhalt and the "Europäischer Fonds für regionale Entwicklung" (EFRE) under Grant No. ZS/2023/12/182057.

References

[1] Wiendahl, H.-P., Wiendahl, H.-H., 2019. Betriebsorganisation für Ingenieure, 9th ed. Hanser, Munich, pp.1421

[2] Bueno, A., Godinho Filho, M., Frank, A.G., 2020. Smart production planning and control in the Industry 4.0 context: A systematic literature review. Computers & Industrial Engineering 149, 106774.

[3] Hopp, W.J., Spearman, M.L., 2000. Factory physics: Foundations of manufacturing management, 2nd ed. Irwin/McGraw-Hill, Boston, pp. 698.

[4] Tanane, B., Bentaha, M.-L., Dafflon, B., Moalla, N., 2025. Bridging the gap between Industry 4.0 and manufacturing SMEs: A framework for an end-to-end Total Manufacturing Quality 4.0's implementation and adoption. Journal of Industrial Information Integration 45, 100833.

[5] Mather, H., Plossl, G., 1978. Priority fixation versus throughput planning. Production and Inventory Management 19 (3), 27–51.

[6] Vollmann, T.E., Berry, W.L., Whybark, D.C., 1997. Manufacturing planning and control systems, 4th ed. Irwin/McGraw-Hill, New York, pp. 836.

[7] Wiendahl, H.-P., 1997. Fertigungsregelung: Logistische Beherrschung von Fertigungsabläufen auf Basis des Trichtermodells, 2nd ed. Hanser, Munich, pp. 382.

[8] Schenk, M., Wirth, S., Müller, E., 2014. Fabrikplanung und Fabrikbetrieb: Methoden für die wandlungsfähige, vernetzte und ressourceneffiziente Fabrik, 2nd ed. Springer, Berlin, Heidelberg, pp. 832.

[9] Schuh, G. (Ed.), 2006. Produktionsplanung und -steuerung: Grundlagen, Gestaltung und Konzepte, 3rd ed. Springer, Berlin, Heidelberg, pp. 876.

[10] Chapman, S.N., 2006. The fundamentals of production planning and control. Pearson/Prentice Hall, Upper Saddle River, New Jersey, pp. 272.

[11] Aamodt, A., Nygård, M., 1995. Different roles and mutual dependencies of data, information, and knowledge - An AI perspective on their integration. Data & Knowledge Engineering 16 (3), 191–222.

[12] Wang, R.Y., Strong, D.M., 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems 12 (4), 5–33.

[13] Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y., 2002. AIMQ: a methodology for information quality assessment. Information & Management 40 (2), 133–146.

[14] Nelson, R.R., Todd, P.A., Wixom, B.H., 2005. Antecedents of Information and System Quality: An Empirical Examination Within the Context of Data Warehousing. Journal of Management Information Systems 21 (4), 199–235.

[15] Sauer, P.C., Seuring, S., 2023. How to conduct systematic literature reviews in management research: a guide in 6 steps and 14 decisions. Review of Managerial Science 17 (5), 1899–1933.

[16] Lingitz, L., Sihn, W., 2020. Concepts to Improve the Quality of Production Plans using Machine Learning. ACTA International Measurement Confederation 9 (1), 32.

[17] Schröer, T., Janßen, J., Schuh, G., Stich, V., 2023. Systematisation Approach: Handling Insufficient Data Quality. Proceedings of the 4th Conference on Production Systems and Logistics, 469–478.

[18] Zong, W., Wu, F., Feng, P.P., 2019. Improving data quality during ERP implementation based on information product map. Enterprise Information Systems 13 (9), 1275–1291.

[19] Messner, M., Pauker, F., Mauthner, G., Frühwirth, T., Mangler, J., 2019. Closed Loop Cycle Time Feedback to Optimize High-Mix / Low-Volume Production Planning. Procedia CIRP 81, 689–694.

[20] Busert, T., Fay, A., 2019. Extended Value Stream Mapping Method: Harmonizing Information Flows for the Control of Production Processes. IFAC-PapersOnLine 52 (13), 54–59.

[21] Reuter, C., Brambring, F., 2016. Improving Data Consistency in Production Control. Procedia CIRP 41, 51–56.

[22] Büscher, C., Voet, H., Krunke, M., Burggräf, P., Meisen, T., Jeschke, S., 2016. Semantic Information Modelling for Factory Planning Projects. Procedia CIRP 41, 478–483.

[23] Denkena, B., Wichmann, M., Arnold, D., 2022. Erfasste Betriebsdaten und ihre Qualität. Zeitschrift für wirtschaftlichen Fabrikbetrieb 117 (12), 847–850.

[24] Ryback, T., Lingitz, L., Gallina, V., Gyulai, D., 2019. Improving the Planning Quality in Production Planning and Control with Machine Learning. Proceedings of the 16th IMEKO TC10 Conference, 131–136.

[25] Soesanto, R.P., Achmad, F., 2024. Leveraging Explainable AI for Accurate Production Forecasting in the Bag Manufacturing Industry. Proceedings of the 7th Asia Pacific Conference on Manufacturing Systems and 6th International Manufacturing Engineering Conference - Volume 2, 105–114.

[26] Dannapfel, M., Wissing, T., Förstmann, R., Burggräf, P., 2019. Human Machine Cooperation in Smart Production: Evaluation of the Organizational Readiness. International Journal of Mechanical Engineering and Robotics Research, 327–332.

[27] Herrmann, J.-P., Tackenberg, S., Padoano, E., Gamber, T., 2021. A literature review and cluster analysis of the Aachen production planning and control model under Industry 4.0. Procedia Computer Science 180, 208–218.

[28] Langer, A., Ortmeier, C., Martin, N.L., Abraham, T., Herrmann, C., 2021. Combining Process Mining And Simulation In Production Planning. Proceedings of the 2nd Conference on Production Systems and Logistics.

[29] Liu, D., Guo, Y., Huang, S., Fang, W., Tian, X., 2023. A stacking denoising auto-encoder with sample weight approach for order remaining completion time prediction in complex discrete manufacturing workshop. International Journal of Production Research 61 (10), 3246–3259.

[30] Urban, M., Koblasa, F., Mendřický, R., 2024. Machine Learning in Small and Medium-Sized Enterprises, Methodology for the Estimation of the Production Time. Applied Sciences 14 (19), 8608.

[31] Vazquez Hernandez, J., Elizondo Rojas, M.D., 2024. Improving spare parts (MRO) inventory management policies after COVID-19 pandemic: a Lean Six Sigma 4.0 project. The TQM Journal 36 (6), 1627–1650.

[32] Lindström, V., Persson, F., Viswanathan, A.P.C., Rajendran, M., 2023. Data quality issues in production planning and control – Linkages to smart PPC. Computers in Industry 147, 1–8.

[33] Arnold, D., Wilmsmeier, S., Denkena, B., Dayeg, A., 2021. Betriebsdaten und ihre aktuellen Potenziale. Zeitschrift für wirtschaftlichen Fabrikbetrieb 116 (11), 852–855.

[34] Myrelid, P., 2017. Information quality deficiencies in delivery schedules and their impact on production scheduling. Production Planning & Control 28 (3), 232–243.

[35] Sagawa, J.K., Nagano, M.S., 2021. Integration, uncertainty and information: how do they affect planning performance? The International Journal of Advanced Manufacturing Technology 28 (1), 23–49.

[36] Schreiber, M., Metternich, J., 2022. Traceability System's Impact On Process Mining in Production. Proceedings of the 3rd Conference on Production Systems and Logistics.

[37] Radke, A.M., Isa, J.b., Tseng, M.M., Reinhart, G., 2019. A Framework for Production Planning and Control as a Service. Proceedings of the 7th IEEE Conference on Systems, Process and Control (ICSPC), 14–19.

[38] Lingitz, L., Gallina, V., Breitschopf, J., Finamore, L., Sihn, W., 2023. Quality in production planning: Definition, quantification and a machine learning based improvement method. Procedia Computer Science 217, 358–365.

[39] Günther, L.C., Colangelo, E., Wiendahl, H.-H., Bauer, C., 2019. Data quality assessment for improved decision-making: a methodology for small and medium-sized enterprises. Procedia Manufacturing 29, 583–591.

[40] Ortmeier, C., Dombrowski, U., 2020. Procedure Model for the Implementation of an Order Release Method. Procedia CIRP 93, 814–819.

[41] de Ruiter, F.J.C.T., Ben-Tal, A., Brekelmans, R.C.M., den Hertog, D., 2017. Robust optimization of uncertain multistage inventory systems with inexact data in decision rules. Computational Management Science 14 (1), 45–66.

Biography

Johannes Jordan has been a researcher and PhD student at the Institute for Engineering of Products and Systems at the Otto von Guericke University Magdeburg since 2025. He is employed by the OpenDANS project, which aims to promote knowledge transfer between science and industry by creating an open platform for testing methods through the exchange of real data. His research interests include production planning and control and factory planning.

Oliver Antons received M.Sc. degrees in both Mathematics and Economics, as well as a doctoral degree in Economics from RWTH Aachen University. He worked as a research associate at RWTH Aachen from 2018 to 2022, and is currently a postdoctoral researcher and group leader for Data-driven Decision Making at Otto von Guericke University Magdeburg. For more information, including his full profile and bibliography, visit https://oliver.antons.eu.

Prof. Dr. Julia Arlinghaus is Chair of Logistics and Supply Chain Management and Director of the Institute for Production and Supply Chain Management at the University of St. Gallen. Her research focuses on new technologies and business models in logistics and manufacturing, methods of artificial intelligence, and new approaches to planning and control for digital and circular value creation systems.

8th Conference on Production Systems and Logistics

Influence Of Legacy Performance Indicators On Contemporary Production And Assembly Research

Tobias Bein[1], Ulf Bergmann[1], Oliver Antons[1], Julia Arlinghaus[2]
[1]*Institute for Engineering of Products and Systems, Otto von Guericke University Magdeburg, Magdeburg, Germany*
[2]*Institute for Production and Supply Chain Management, University St. Gallen, St. Gallen, Switzerland*

Abstract

Scientific credibility is essential for both researchers and practitioners alike. Irreproducible empirical research or questionable citation practices lead to problems regarding transparency or simply hinder scientific discourse and knowledge gain. Especially regarding frequently used performance indicators, such as production cost or lead time, industry and academics need credible data alike for their strategic orientation, like investment decisions or the derivation of research gaps and hypothesis. This paper examines citation patterns of fundamental production and assembly indicators. In particular, the key performance indicators of assembly cost and assembly time are analyzed twofold, based on a highly cited German textbook on assembly. First, a citation analysis that looks back at past research is performed to show the origins of the assembly performance indicators. Second, a citation content analysis looking forward is carried out on literature citing the German assembly textbook to provide an overview of the relevance of the examined performance indicators. Lastly, the findings are discussed and implications for scientific credibility in assembly and production research are derived. The results demonstrate that the origins of the exemplary assembly performance indicators are questionable and difficult to determine, as older literature becomes distorted, is inaccessible or fails to give credit to previous work. Furthermore, the second literature review reveals that, despite the indistinct origins of the assembly performance indicators, they remain highly cited. With the indicators being used for motivating many authors' works or as state of the art, an adverse impact on scientific credibility is identifiable.

Keywords

Manufacturing Systems; Assembly Systems; Assembly Cost; Assembly Time; Key Performance Indicators

1. Introduction

"Standing on the shoulders of giants" is a term often associated with Sir Isaac Newton, describing how one's achievements are often build upon preceding works of other people. In the scientific context this is associated with using and expanding theoretical models, developed methodologies or empirical insights. Citations perform this important role in research, functioning as part of a scientific reward system [1]. However, authors like Anderson and Lemken [2] or Small [3] describe current citation practices as "far from optimal". For example, Anderson and Lemken [4] examine that many citations are only peripheral, not adding substance to the citing authors' arguments, titling them as throw-away references. This is also influenced by authors citing widely known or highly-cited works to avoid making it apparent that they do not comprehend the theory in question [5]. Citing widely known works is not only convenient for authors, but also an issue for reviewers who have to verify if the cited work supports the citing works assertions [2].

DOI: https://doi.org/10.15488/21003
ISSN: 2701-6277

For the analysis of production and assembly systems, academia and industry utilize key performance indicators, which are quantifiable and strategic measures used to understand and improve system performance by minimizing or maximizing the indicators [6]. Typical indicators used for decision-making in production systems revolve around cost, time, quality and flexibility [7]. Such indicators are often surveyed in empirical analysis or developed through the utilization of benchmark datasets. An example of the latter are the abstract scheduling datasets by Taillard [8], which are, however, restrictive in terms of practical applicability due to their deterministic machining time and missing data typically found in industrial companies. This leads to a strict dependency on the context to effectively utilize such indicators, for which correct referencing is significant. Without referring and considering the context, researchers may derive wrong assumptions and research needs, limiting a works knowledge gain. For industry, indicators are often used as benchmarks for similar systems [9,10], which can lead to suboptimal decision-making and poor performance if the chosen indicators themselves are unfitting [11].

The key measures motivating our work are assembly cost and time, for which the same indicators can be found in recent research (e.g. [12,13]), as well as textbooks dedicated to industrial trainees and students (e.g. [14]). For the indicators in question, the context is often generalized for production and assembly as a whole and many works are missing references completely. There are also indications that the indicators themselves are several decades old [15,16]. Besides the non-optimal citation practice, the generalized context and age of both indicators can lead to the previously described adverse influences on academia and industry, which is why an analysis of these indicators firmness in contemporary times is required.

In regard to the highlighted problems within citation practices and indicator usage in production and assembly research, this paper investigates how influential these fundamental and widely known production and assembly indicators are in recent research. Since this cannot be answered holistically, we concentrate our efforts on the introductory chapter of a German textbook by Lotter [15,16] as the basis for our analysis, for which we have identified that it contains both cost and time indicators. The selection is also based on the fact that the work itself is highly cited, from which we expect that both indicators are referenced by subsequent publications, mainly because the indicators of cost and time are relevant and applicable to many topics in, e.g., assembly design or automation. Because the textbook is a secondary source, we first unravel the citation patterns towards the original empirical studies by tracing back any available reference of Lotter's work. This method is named historic citation analysis in the following, and should provide insights into the original sources, their contents and contexts. Secondly, a citation content analysis sheds light on the relevance and utilization of the assembly indicators today. This includes their critical assessment by authors nowadays. By following this approach, we aim to highlight the importance of scientific work, especially when dealing with indicators from secondary literature.

The remainder of this paper is structured in four sections. The following section details the need for credible key performance indicators in production and assembly and presents the assembly performance indicators used for the analysis. Section 3 details the methodical approach for answering the research questions. In Section 4, we report on and discuss the analysis results, and Section 5 concludes our paper.

2. Theoretical Background

2.1 The Need for Credible Performance Indicators

To achieve credible science and practices, Taylor et al. describe them as being verifiable for others to build on previous work in academia and in industry [17]. With manufacturing paradigms like Industry 4.0 and increasingly data-driven digital manufacturing, research and industry are able to enhance the calculation of benchmark indicators, which can be used to compare similar operations between industry enterprises, if the indicators are standardized [9]. However, if deficient indicators are used, the original purpose of performance indicators enabling a management by facts to identify and eliminate problems [18], and to be the basis to

facilitate decision-making [19], is impaired. This leads to suboptimizations, conflicts and malpractices [20]. In these situations, indicators display false or manipulated realities [20], resulting in wrong conclusions and hypothesis being drawn, directly effecting subsequent work. Therefore, knowledge on the content and context of any indicator are essential when utilizing them in research or industry alike.

2.2 Assembly Cost

Assembly consists of all operations for joining, handling, checking and adjusting geometrically determined parts and components [16]. In the last decades, many of these operations were automated, but manual assembly is still widely spread in many companies according to recent surveys [21–23]. Especially with a high proportion manual labor, assembly is a cost intensive process, for which rationalization measures are important. We therefore analyze the production and assembly cost indicators from Lotter's textbook [15,16] in our study, which are illustrated in Figure 1.

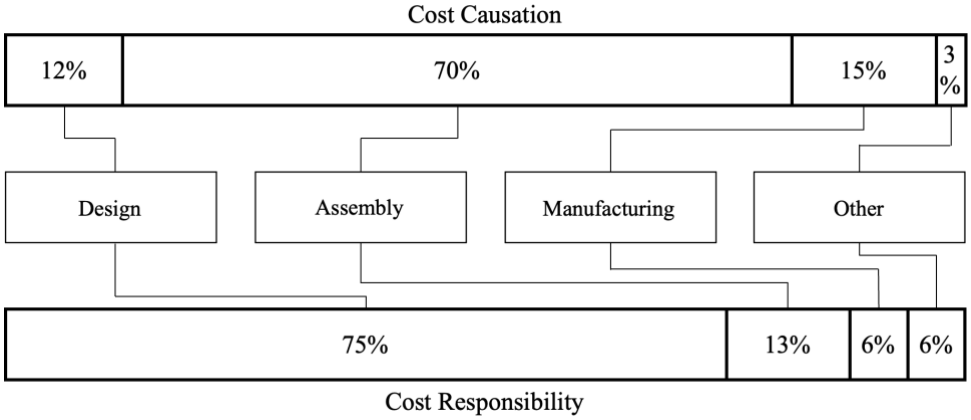


Figure 1: Cost Causation and Responsibility in Assembly, based on [16,24]

The assembly cost indicators consist of the cost causation and cost responsibility of product design, assembly, manufacturing and other production related areas, which can be purchasing. Cost causation comprises costs incurred during the actual production. On the other hand, cost responsibility describes the extent to which departments influence the costs incurred during production. For assembly in particular it can be seen that the percentage of production cost responsibility is relatively small with 13%, compared to the 70% of cost causation. In contrast, product design accounts for only 12% of cost causation, but is responsible for 75% of costs. However, the cost indicators in question are linked to a work from 1981 [24], which itself is not the original source.

2.3 Assembly Time

An important aspect of the value creation process is the time it takes to produce a good, which splits up into the time for parts manufacturing and assembly. The attributes of cost and time are interrelated [7], with both indicators effecting each other, which is why it is selected as the second indicator for our study. Ranges of this indicator are illustrated in Lotter's textbook [15,16] for different industries. While the percentage of general assembly is somewhere between 15-70%, the individual industry branches are situated in between. The assembly in the mechanical engineering industry accounts for 20-45% and in automotive 30-50% of overall process time. The highest percentage of assembly time is present in electrical and precision mechanics with 40-70%. These assembly times are referenced towards a previous work from 1992 [25], which is also only a secondary source.

3. Methodic Approach

Since this papers goal is to analyze publications from a textbook in a detailed manner, we adapt a twofold methodic approach. To identify the foundations of the empirical research and to answer how their content

and context have changed over the years, we use a historic citation analysis approach. To identify recent literature and its handling of the assembly indicators, a citation content analysis is performed.

In the historic citation analysis, we trace back every mention of the assembly indicators towards their origin. The concept of this approach is similar to a citation content analysis as an historical tool [2], but going backwards in time. In doing so we identify the citation patterns needed to trace the indicators as well as their origin itself. With the citation patterns it becomes possible to track any changes in context or content, such as combining the knowledge of previous authors. In order to illustrate the broader context of the assembly indicators, we perform the historic citation analysis in an iterative manner. We start at the textbook works of Lotter [15,16] in the first iteration. We then perform the citation content analysis explained in the following, through which it becomes possible to build a "citation image" [26], which describes a network of important works regarding a topic [2]. These additional works allow us to supplement our historic citation analysis in order to illustrate a broader citation pattern.

The second analysis follows the methodology of a citation content analysis as described by Anderson and Lemken [2]. We therefore gather all citations to the two editions of Lotter's textbook [15,16] to answer the research question stated beforehand. To identify citing works, we use Google Scholar's and Semantic Scholar's cited-by feature. While Semantic Scholar differentiates between the editions of the textbook, Google Scholar does not. We retrieved 624 papers from Google Scholar, and from Semantic Scholar 202 and 101 publications for the editions, respectively. The 927 works were screened, whereby 270 duplicates were first removed. In the filtering process, 8 works were excluded for their language not being German or English, which are the languages the authors are fluent in. 25 papers were not included due to being non-identifiable, mostly through inconclusive titles or authors. 72 works could not be retrieved. We therefore assessed 552 full-text papers for eligibility, from which 5 did not include a reference to the textbook and 461 did not mention the assembly indicators in question. From the remaining data 4 publications used the assembly indicators, but referenced other works identified through the historic citation analysis, with the works being included in Figure 3 or Figure 4. As the indicator origins and meanings stay the same, they are included in the final sample. The remaining 86 publications are shown in Figure 2.

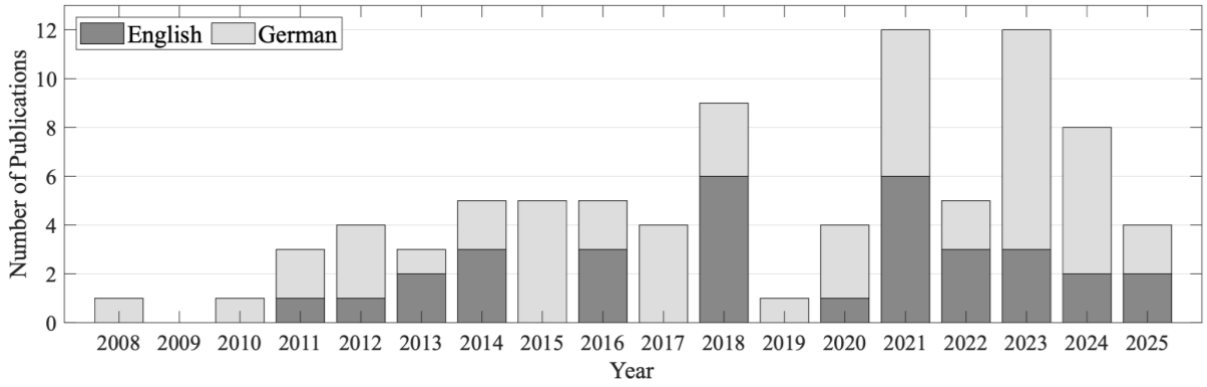


Figure 2: Data Sample for the Citation Content Analysis, sorted by Year of Publication and Language

Because assessing the substance on citation contexts goes beyond counting, we identify the context and purpose for which the assembly indicators are cited. We first analyze objective characteristics, such as the document type and which assembly indicator is being cited. A distinction is then made on the overall topic of the citing work to better categorize the data sample. This categorization is developed through iterative discussion rounds between the authors. As mentioned before, it is important that publications critically assess and deal with the knowledge claims of works they are citing. This aspect is analyzed by reviewing if publications either critically discuss the assembly indicators or use multiple sources to strengthen their argumentation.

4. Results and Discussion

4.1 Historic Citation Analysis

To identify the origins of the assembly indicators for cost responsibility and cost causation, as well as the assembly time, we started at the textbook chapter of Lotter [15,16] and trace back every citation mentioned in connection with the assembly indicators. Because many works are not digitalized, an intensive amount of time was invested into researching and ordering the works from university libraries across Germany. The resulting citation patterns for tracing back the assembly indicators of cost and time are illustrated in Figure 3 and Figure 4. Every white colored box symbolizes the first iteration of our historic analysis, placing an emphasis on works traceable through Lotter's textbook. The gray colored rectangles showcase the works included through the citation image, which are co-cited publications identified through the citation content analysis presented in the subsequent Section 4.2. The figures also include works without specific mentions of the indicators and works which we could not access, demonstrating their role in the citation pattern. This is believed by the authors to help in understanding the approach used and to support transparency. Other depictions of the assembly indicators, which have been identified through the citation image, are illustrated in the appendix.

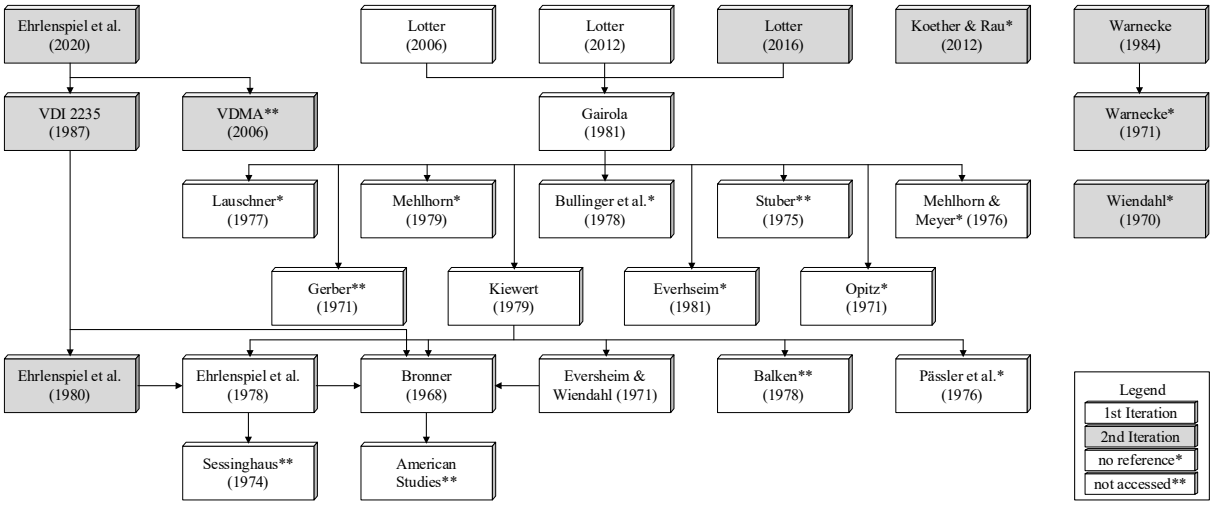


Figure 3: Citation Pattern of Assembly Cost, based on [15,16,27–50]

For the assembly cost indicators, it can first be seen that Lotter's textbooks [15,16] cite Gairola [24], who summarizes the parameters for the cost causation and responsibility, resulting in the illustration depicted in Figure 1. Many of the following preceding sources, however, do not indicate if their claim is based on own experiences, studies or previous work. Some works do not even mention assembly cost explicitly, but provide information on time indicators instead [34,41], with a conclusion that much time needed equals high cost. For assembly cost causation, the accessible works, even the specifically mentioned one [38], did not provide insights on their origin. Kiewert [36] is specifically cited for cost responsibility, and is critically assessing the topic in detail. The 75% of costs the design is responsible for can only be meant for unnecessary or extra costs atop the ideal production costs, because the latter cannot be reduced [36]. Kiewert further argues that designers are causing these unnecessary costs rather than being responsible for them, because responsibility implies monitoring the effects of one's own actions, which would require them to be knowledgeable on costs incurred in subsequent production steps [36]. Further references point towards the works of Bronner [28] and Ehrlenspiel et al. [31], which center around the value engineering method. While Ehrlenspiel et al. [31] themselves refer to a student thesis [44] and Bronner [28], the latter points towards unnamed American studies for all cost responsibility indicators. While Bronner utilizes the same data for cost responsibility used in Figure 1, he labels 13% being work preparation rather than assembly. It can therefore be assumed that the label of assembly in Figure 1 may be distorted through different authors over time. Through the citation image we could identify further works mentioning the cost responsibility and

causation [32,37,46,48,49], as well as a report on an empirical survey [30]. Ehrlenspiel et al. [30] questioned 42 companies from mechanical engineering, automotive and precision engineering, most of them having more than 500 employees, on cost reductions through value engineering. Their analysis determines that for 33% of unnecessary costs identified by the value engineering method, the responsibilities are 65% on design, 19% on work preparation and 15% on procurement. Ehrlenspiel et al. [30] therefore consider Bronner's [28] findings to be verified. The results of Bronner [28] and Ehrlenspiel et al. [30] are later used in the VDI 2235 [46] and Ehrlenspiel et al. [32], both also not labeling any assembly indicator. Since all these works refer to work preparation rather than assembly, this reinforces our assumption that the label in Gairola [24] and consequently Lotter's work [15,16] has been distorted. The citation image could not answer the question on the origins of the cost causation. Although the textbook by Koether and Rau [37] states that assembly causes 20-40% or more of overall production costs, it does not specify how these indicators came to be. Many works [46,51,52], some even in the early 1970s [49,50], often refer to an unspecified work of Bronner and Holste, which depicts cost responsibility and includes data for the cost causation. However, in 1970, Wiendahl [50] comments that the data shown is qualitative in nature. It also has to be mentioned that the cost causation depicted in these works and the one's from Gairola [24] and Lotter [15,16] differ substantially. For the assembly cost indicators in Lotter's textbook, it can therefore be seen that their origins cannot be clearly identified, but with some of the contents dating back nearly 60 years and seemingly being distorted.

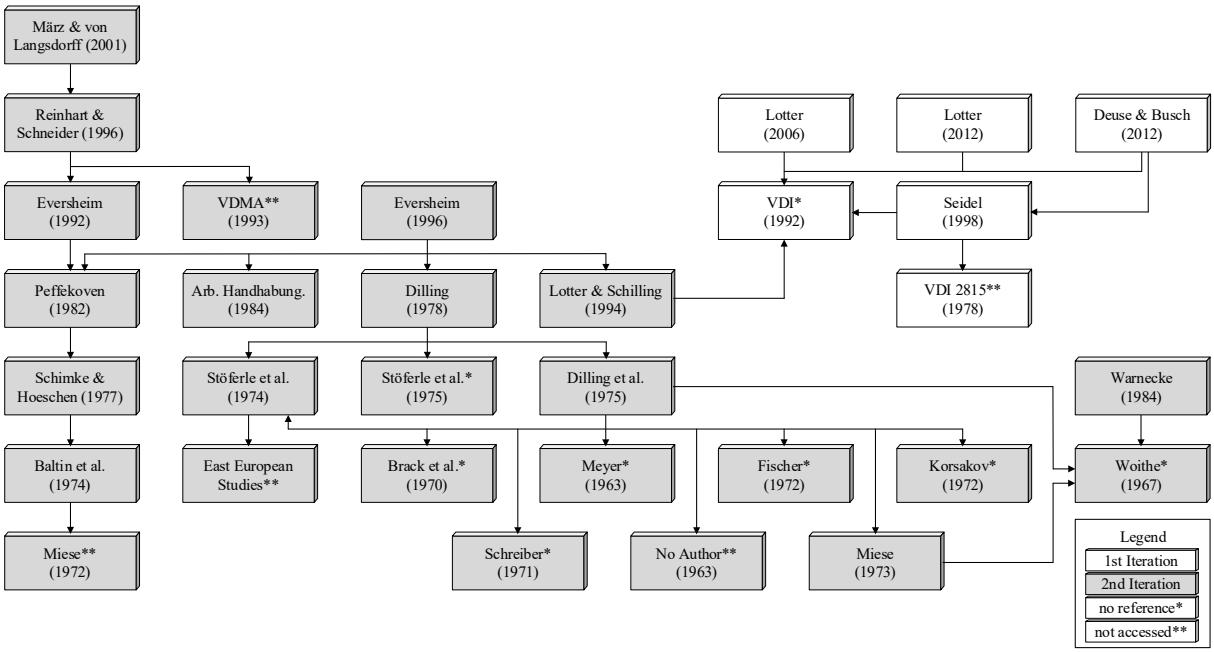


Figure 4: Citation Pattern of Assembly Time, based on [15,16,25,48,51,53–77]

For the assembly time indicators, the different chapters in Lotter's textbook mentioning them refer to a previous work of the VDI [25], which mentions the information in the introductory chapter, but with other data than the citing work. While Lotter's work [15,16] defines 15-70% for general assembly time and 20-45% assembly time in mechanical engineering, the VDI [25] refers to 20-70% general assembly time and up to 25% assembly time in the mechanical engineering industry. Because no further reference is given, we identified the work of Eversheim [51] as the starting point for the second iteration of our historic citation analysis, because the work mentions the same industries with the same or similar data. Peffekoven [68] describes the assembly indicators of different industries and concludes a mean assembly time of 25-35%. The author details 40-70% for assembling machine tools, 5-10% for gearbox assembly and 25-45% for engine assembly. Even though the work does not indicate how these indicators were surveyed and if the work is the original source, subsequent work [54,70] in the citation pattern reveals much less information on the content in question and eventually refers to an unpublished industry study [65]. The results of a large-scale empirical study in the German assembly industry with a focus on assembly automation and robots are

presented in [53], highlighting the industries of mechanical engineering, automotive, electric and precision mechanics. The works study design consists of a questionnaire with 355 respondents, 43 case studies and several expert interviews specifically for low to high volume assembly. Even though the study highlights the distributions of takt times, rationalization measures for planned assembly times and the production life cycle times in the different industries, it does not explicitly mention any ranges for the assembly time as part of the total production time. It does however survey that assembly causes 20-50% of total production cost [53]. Following the citation pattern, the work of Dilling et al. [58] reviews different publications on assembly time and concludes that many authors (e.g. [61,64,71]) estimate it to be around 40-60% of the total manufacturing time. Detailed analysis can be found in Brack et al. [55] and Woithe [77]. The latter describes studies analyzing the assembly time in 11 different industries in the GDR, ranging from 15-75% with a mean of 40-45% [77]. Brack et al. [55] describe a range for different industries between 7.5% and 85%, with a mean of 36.3%. The authors however only sum up an undisclosed previous study from the CSSR. From the different publications, the assembly time indicators in Lotter's textbook seem to be derived from the large-scale assembly survey in the German industry [53], because the same industries are reviewed. However, the origins of the data used in VDI [25] and subsequently in Lotter [15,16] could not be identified.

The historic citation analysis of both assembly indicators reveals that both indicators are a partially distorted and untraceable legacy of academic work, which often precedes the 1980s and partly the 1970s. The authors therefore argue that both indicators should be seen as legacy measurements and should not be used by contemporary work describing current trends or a state of the art in production and assembly.

4.2 Citation Content Analysis

The citation content analysis focuses on the 86 identified works citing Lotter's publication and using the assembly indicators. To better discern the contents and contexts, a distinction is first made based on the works topics, the cited assembly indicators, the document types and the way the indicators are cited and critically assessed. An overview of the results is illustrated in Figure 5.

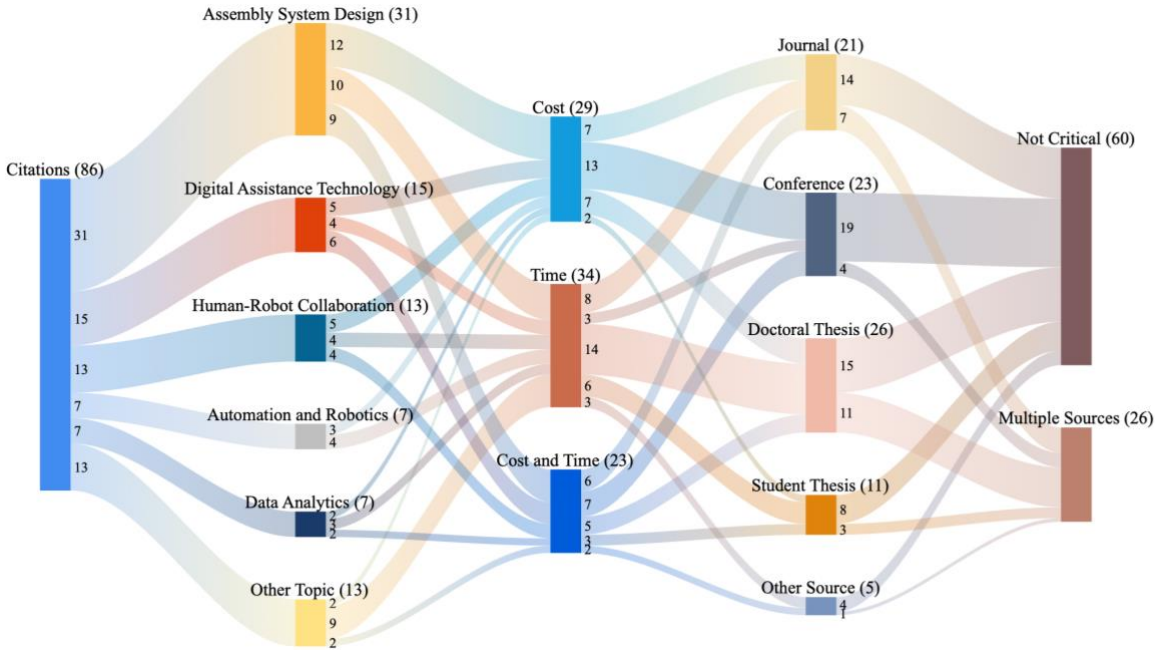


Figure 5: Results of the Citation Content Analysis. The Sankey Plot depicts the Publications Topic, cited Assembly Indicator, Document Type and the Authors critical Evaluation of the Indicator's Knowledge Claim

It can first be seen that most of the publications are categorized as assembly system design, followed by digital assistance technologies and human-robot collaboration. The least amount of works belong to the topics of automation and robotics, as well as data analytics. The category 'Other Topic' consists of publications in the areas of procurement, product design or workforce management, which did not fit into

any of the previous topics. The document types most prominently citing the assembly indicators are doctoral thesis in assembly design and the miscellaneous category, followed by journals and conference papers. In 63% of all works the assembly indicators are utilized in the introduction, often as a form of motivation to depict current problems related to assembly systems. The others works cite the indicators in the theoretical background or as state of the art. Figure 5 depicts that each indicator is being cited fairly equally, with many works referencing both assembly cost and time. In detail, the most mentioned data are the 70% of cost causation by assembly and the general assembly time of 15-70%. Regarding the critical assessment of Lotter's textbooks knowledge claims, 70% of analyzed works simply cite both indicators without further evaluation. Another 30% cite further work to reinforce the assembly indicators. Not one single publication reviewed questions the novelty, origin or the validity of the indicators in their own work.

The results demonstrate that despite unclear origins and distorted content, both indicators remain highly cited in recent works. This can implicate the authors assumptions and hypothesis, and possibly the expected contribution value of their research. Especially with the indicators being used in works dealing with real-world applications of research, they can lead to suboptimizations through misleading interpretations of results [20]. Overall, this also affects the credibility of production assembly science, with many publications utilizing seemingly outdated performance measures. This is not least because citation practices are not optimal in the exemplary cases analyzed, with authors not critically reflecting previous knowledge claims. The latter seems to be especially important, as many works are doctoral or student thesis, demonstrating that these debatable indicators are already being taught to or used by the next generation of academics and practitioners. However, the results did not include any specific mention of these indicators being used as benchmarks by industrial companies.

4.3 Limitations of the Study

There are several limitations to our study, the first being that we are only examining two specific assembly indicators, which are not holistically representative for the production and assembly literature. Another limitation is the textbook [15,16], which has been used as a starting point for both citation analysis approaches. The work limits the scope for the historic citation analysis as well as the citation content analysis, because using more works from the same author or more works in general may lead to a bigger citation pattern from which better conclusions on the indicators can be derived. Despite a detailed and time intensive research, both citation analysis methods are inconvenienced by some of the works not being accessible for our study, which may contain crucial information for our analysis. Even though many publications have been analyzed in the citation content analysis, there exist works mentioning the same indicators, including the specific numbers and industries, but without citing their origin [13,14]. We also showcased many other works utilizing the indicators, which themselves have been frequently cited. We therefore expect a number of unrepresented novel publications to be present, which cannot be analyzed in this study.

5. Conclusion

Within our study, we have critically examined the exemplary performance indicators of assembly cost and time as presented in a widely cited textbook by Lotter [15,16]. The theoretical contribution of our work is that both indicators display data which is several decades old and can either not be traced to an empirical source or include contents that have been distorted over time. This indicates a gap between research and practice with regard to the data for each indicator. Authors, especially the interest groups of assembly system design, human-robot collaboration and digital assistance technology highlighted in Section 4.2, should not use both indicators as means for motivating their work or demonstrating the state of the art in assembly without a thorough critical evaluation. From a practical perspective, industry should avoid using these indicators as benchmarks in their companies. Consequently, with a consistent interest in both indicators, future work should qualitatively or quantitatively survey different assembly industries to get a novel overview on both assembly cost and time, as well as other related performance measures.

Acknowledgements

This research was supported by the state of Saxony-Anhalt and the "Europäischer Fonds für regionale Entwicklung" (EFRE) under Grant No. ZS/2023/12/182057.

Appendix

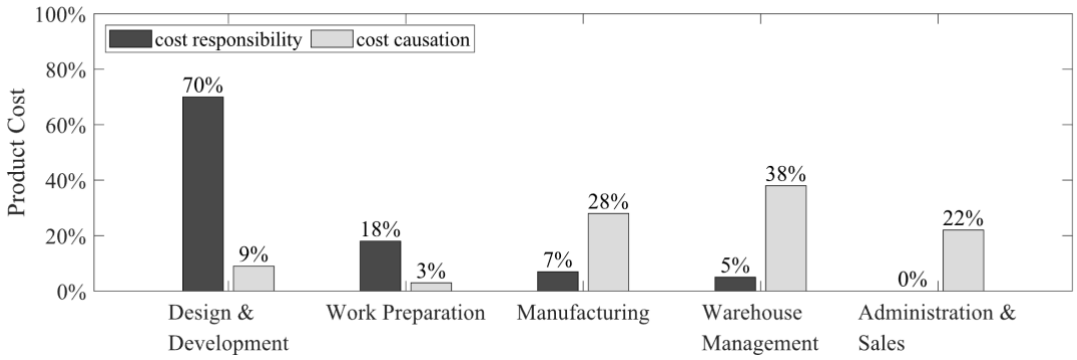


Figure 6: Cost Causation and Cost Responsibility in different Company Divisions, based on Ehrlenspiel et al. [32], who reference VDI 2235 [46] and VDMA [47]

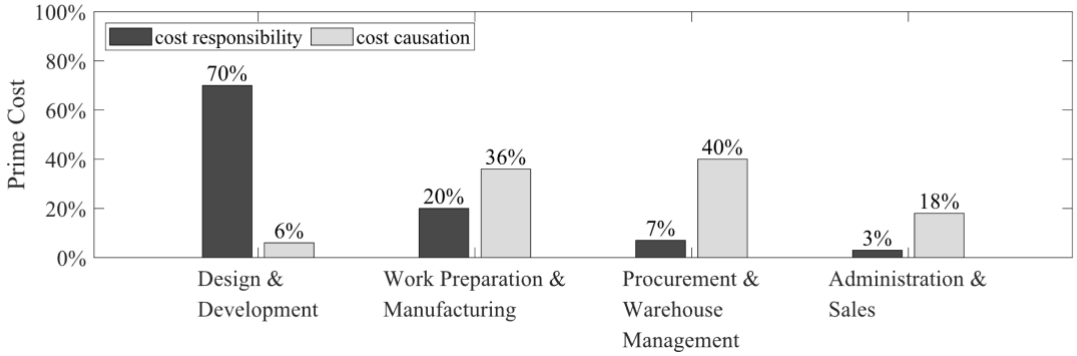


Figure 7: Cost Responsibility and Cost Causation in different Company Divisions, based on VDI 2235 [46], which references Bronner [28] and Ehrlenspiel et al. [30]

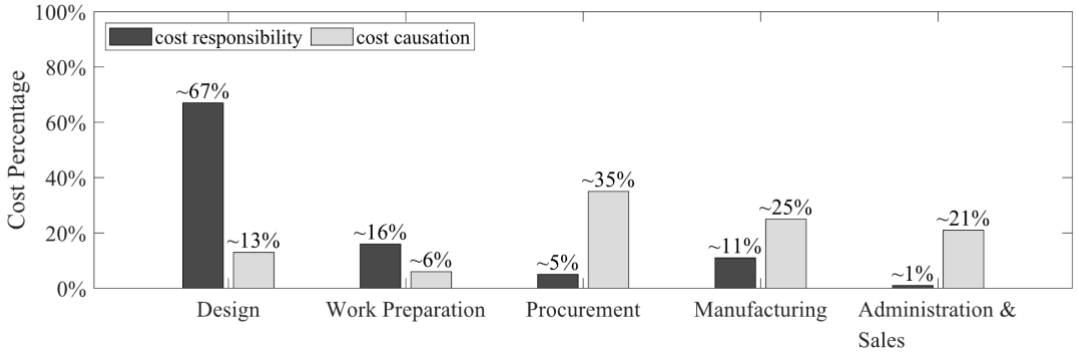


Figure 8: Cost Responsibility and Cost Causation in different Company Divisions, derived from Illustrations in [51,52] referring to Bronner and Holste's Work

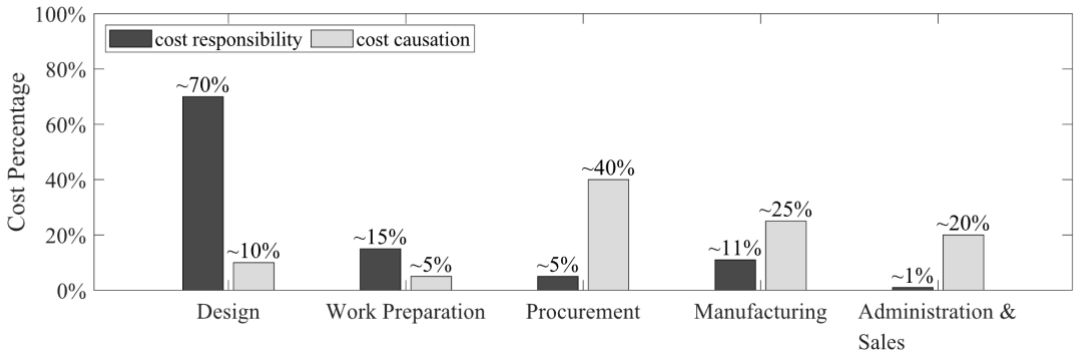


Figure 9: Cost Responsibility and Cost Causation in different Company Divisions, derived from Illustrations in [49,50] referring to Bronner and Holste's Work

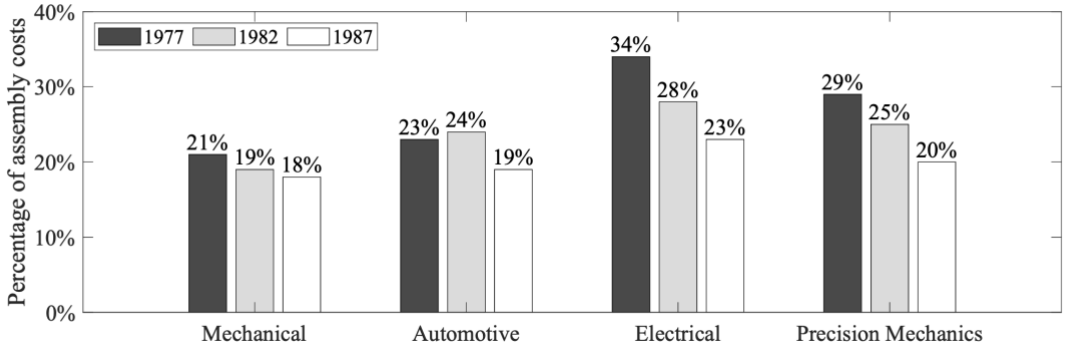


Figure 10: Mean Percentage of Assembly Cost Causation in Total Production Cost for different Industries, based on [53]. The individual Values range between 20-50%

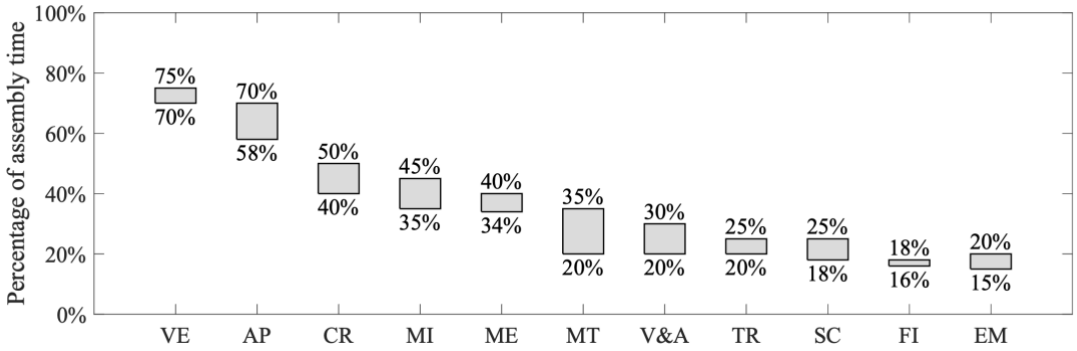


Figure 11: Percentage of Assembly Time in Total Production Time for different Industries, based on [77]. The mean Assembly Time over all Industries is 40-45%. VE = Vessels, AP = Apparatuses, CR = Cranes, MI = Measurement Instruments, ME = Metallurgical Equipment, MT = Machine Tools, V&A = Vehicle & Agricultural Machinery, TR = Transmissions, SC = Steel Construction, FI = Fittings, EM = Electric Motors

References

[1] Merton, R.K., 1957. Priorities in Scientific Discovery: A Chapter in the Sociology of Science. American Sociological Review 22 (6W), 635–659.

[2] Anderson, M.H., Lemken, R.K., 2023. Citation Context Analysis as a Method for Conducting Rigorous and Impactful Literature Reviews. Organizational Research Methods 26 (1), 77–106.

[3] Small, H., 1982. Citation context analysis, in: Dervin, J., Voigt, J.M. (Eds.), Progress in communication science, vol. 3. Ablex, pp. 287–310.

[4] Anderson, M.H., Lemken, R.K., 2019. An Empirical Assessment of the Influence of March and Simon's Organizations : The Realized Contribution and Unfulfilled Promise of a Masterpiece. Journal of Management Studies 56 (8), 1537–1569.

[5] Sutton, R.I., Straw, B.M., 1995. What Theory is Not. Administrative Science Quarterly 40 (3), 371–384.

[6] International Organization for Standardization (ISO), 2014. ISO 22400-2:2014: Automation systems and integration - Key performance indicators (KPIs) for manufacturing operations management: Part 2: Definitions and descriptions 25.040.01.

[7] Chryssolouris, G., 2006. Manufacturing systems: Theory and practice, 2nd ed. Springer, Berlin, New York, Heidelberg.

[8] Taillard, E., 1993. Benchmarks for basic scheduling problems. European Journal of Operational Research 64 (2), 278–285.

[9] International Organization for Standardization (ISO), 2014. ISO 22400-1:2014: Automation systems and integration - Key performance indicators (KPIs) for manufacturing operations management: Part 1: Overview, concepts and terminology 25.040.01.

[10] Lindberg, C.-F., Tan, S., Yan, J., Starfelt, F., 2015. Key Performance Indicators Improve Industrial Performance. Energy Procedia 75, 1785–1790.

[11] Schrage, M., Muttreja, V., Kwan, A., 2022. How the wrong KPIs doom digital transformation. MIT Sloan Management Review 63 (3), 35–40.

[12] Buckhorst, A.F., Grahn, L., Schmitt, R.H., 2022. Decentralized Holonic Control System Model for Line-less Mobile Assembly Systems. Robotics and Computer-Integrated Manufacturing 75, 102301.

[13] Refflinghaus, R., Kern, C., 2019. Analyse der menschlichen Zuverlässigkeit für manuelle Montagetätigkeiten, in: Sträter, O. (Ed.), Risikofaktor Mensch? Zuverlässiges Handeln gestalten. Beuth, Berlin, Wien, Zürich, pp. 191–211.

[14] Schmid, D., Koke, T., Konold, P., Kaufmann, H., Maier, M., Kirchner, A., 2018. Industrielle Montage. EUROPA-LEHRMITTEL, Haan-Gruiten.

[15] Lotter, B., 2006. Einführung, in: Lotter, B., Wiendahl, H.-P. (Eds.), Montage in der industriellen Produktion. Ein Handbuch für die Praxis. Springer, Berlin, New York.

[16] Lotter, B., 2012. Einführung, in: Lotter, B., Wiendahl, H.-P. (Eds.), Montage in der industriellen Produktion. Springer, Berlin, Heidelberg.

[17] Taylor, S.J., Eldabi, T., Monks, T., Rabe, R., Uhrmacher, A.M., 2018. Crisis, What Crisis - Does Reproducibility in Modeling & Simulation Really Matter? Proceedings of the 2018 Winter Simulation Conference, 749-762.

[18] Kaganski, S., Majak, J., Karjust, K., Toompalu, S., 2017. Implementation of Key Performance Indicators Selection Model as Part of the Enterprise Analysis Model. Procedia CIRP 63, 283–288.

[19] Mourtzis, D., Fotia, S., Doukas, M., 2015. Performance Indicators for the Evaluation of Product-Service Systems Design: A Review, in: Umeda, S., Nakano, M., Mizuyama, H., Hibino, H., Kiritsis, D., Cieminski, G. von (Eds.), Advances in Production Management Systems: Innovative Production Management Towards Sustainable Growth. IFIP WG 5.7 International Conference, APMS 2015, Tokyo, Japan, September 7-9, 2015, Proceedings, Part II, vol. 460. Springer International Publishing, Cham, pp. 592–601.

[20] Brenner, J., 2019. Shopfloor Management und seine digitale Transformation: Die besten Werkzeuge in 45 Beispielen: mit 116 Bildern und 2 Tabellen. Hanser, München.

[21] Bank, L., Luber, M., Theumer, P., Zipfel, A., Kämpfer, T., Hiller, T., Heuer, T., Demke, T.M., Mundt, C., Köster, N., Janke, T., Maibaum, J., Schmidthuber, M., 2021. PPS-Report 2021: Studienergebnisse.

[22] Fulterer, J., Maetschke, J., Kuhn, C., Kalchschmid, V., Karg, F., Luber, M., Demke, T.M., Hiller, T., Kämpfer, T., Steffens, C., Mundt, C., Alieksieiev, V., 2023. PPS-Report 2023: Studienergebnisse.

[23] König, M., Winkler, H., 2023. Digitalisierung in der Montage mit dem Schwerpunkt des Einsatzes von Assistenzsystemen.

[24] Gairola, A., 1981. Montagegerechtes Konstruieren Ein Beitrag zur Konstruktionsmethodik. Dissertation, Darmstadt.

[25] VDI (Ed.), 1992. Flexible Montage. VDI-Verlag, Düsseldorf.

[26] Furner, J., 2003. Little Book, Big Book. Journal of Librarianship and Information Science 35 (3), 189–201.

[27] Balken, J., 1978. Auswertung von Wertanalysen zur Ermittlung von Kosteneinflüssen und Hilfsmitteln zum kostenarmen Konstruieren. Technische Universität München. Lehrstuhl für Konstruktion i.M., DFG Abschlußbericht, München.

[28] Bronner, A., 1968. Wertanalyse als integrierte Rationalisierung. Werkstattstechnik 58 (1), 16–21.

[29] Bullinger, H., Emele, R., Bauch, H., 1978. Informationsverarbeitung analysiert: Die Produktivitätssteigerung fängt bereits in der Konstruktion an. VDI-Nachrichten 41, October 13, 9.

[30] Ehrlenspiel, K., 1980. Möglichkeiten zum Senken der Produktkosten - Erkenntnisse aus einer Auswertung von Wertanalysen. Konstruktion 32 (5), 173–178.

[31] Ehrlenspiel, K., Kiewert, A., Lindemann, U., 1978. Produktionskosten senken - eine Aufgabe der Konstruktion. Konstruktion 30 (4), 149–154.

[32] Ehrlenspiel, K., Kiewert, A., Lindemann, U., Mörtl, M., 2020. Kostengünstig Entwickeln und Konstruieren, 8th ed. Springer, Berlin, Heidelberg.

[33] Eversheim, W., Wiendahl, H., 1971. Rationelle Auftragsabwicklung im Konstruktionsbereich. Giradet, Essen.

[34] Eversheim, W., 1981. Organisation in der Produktionstechnik 1. Grundlagen. VDI-Verlag, Düsseldorf.

[35] Gerber, H., 1971. Forderungen an die Konstruktion. Ein Beitrag zur Konstruktionsmethodik. Doktorarbeit, Braunschweig.

[36] Kiewert, A., 1979. Systematische Erarbeitung von Hilfsmitteln zum kostenarmen Konstruieren. Dissertation, München.

[37] Koether, R., Rau, W., 2012. Fertigungstechnik für Wirtschaftsingenieure, 4th ed. Hanser, München.

[38] Lauschner, H.J., 1977. Systematische Anwendung von Kunststoffen bei der Realisierung technischer Funktionselemente und -gruppen. Konstruktion 29 (10), 408–415.

[39] Lotter, B., Deuse, J., Lotter, E., 2016. Die Primäre Produktion. Springer Berlin Heidelberg, Berlin, Heidelberg.

[40] Mehlhorn, H., 1979. Zu einigen fertigungstechnischen Aspekten in der Gerätetechnik. Feingerätetechnik 28 (10), 456–461.

[41] Mehlhorn, H., Meyer, B., 1976. Mechanisierung und Automatisierung von Fügeprozessen für Kleinteile. Feingerätetechnik 25 (11), 494–496.

[42] Opitz, H. (Ed.), 1971. Produktplanung - Konstruktion - Arbeitsvorbereitung. Giradet, Essen.

[43] Pässler, E., Eckhardt, W., 1976. Grundfragen der Rationalisierung der Arbeit des Konstruktionsingenieurs. Maschinenbautechnik 25 (5), 197–201.

[44] Sessinghaus, P., 1974. Methoden der kreativen Ideenfindung bei der Wertanalyse. Abschlußarbeit, Mönchengladbach.

[45] Stuber, P., 1975. Parts transfer and automatic equipment. Tooling and Produktion.

[46] VDI, 1987. VDI-Richtlinie 2235:1987-10: Wirtschaftliche Entscheidungen beim Konstruieren; Methoden und Hilfen. VDI Verlag.

[47] VDMA, 2006. Kennzahlenkompass 2002. Maschinenbau-Verlag, Frankfurt Main.

[48] Warnecke, H.J., 1984. Der Produktionsbetrieb: Eine Industriebetriebslehre für Ingenieure. Springer, Berlin, Heidelberg.

[49] Warnecke, H.-J., 1971. Einflüsse und Tendenzen in der Produktionstechnik. wt - Zeitschrift industrielle Fertigung 61 (10), 667–671.

[50] Wiendahl, H.-P., 1970. Funktionsbetrachtungen technischer Gebilde: Ein Hilfsmittel zur Auftragsabwicklung in der Maschinenbauindustrie. Dissertation.

[51] Eversheim, W., 1996. Organisation in der Produktionstechnik 1: Grundlagen, 3rd ed. Springer, Berlin, Heidelberg.

[52] Westkämper, E., 2006. Einführung in die Organisation der Produktion. Springer, Berlin, Heidelberg.

[53] Arbeitsgemeinschaft Handhabungssysteme, 1984. Einsatzmöglichkeiten von flexibel automatisierten Montagesystemen in der industriellen Produktion: Montagestudie. VDI, Düsseldorf.

[54] Baltin, E., Chondros, G., Hölken, W., Kirmse, W., Miese, M., Schimke, E., Schuler, L., 1974. Systematische Planung und Abwicklung der Montage. Industrie-Anzeiger 96 (70), 1604–1611.

[55] Brack, G., Klitzsch, F., Piegert, R., 1970. Automatisierung im Maschinenbau: Grundlagen, Gerätetechnik, technologische Prozesse. VEB Verlag Technik, Berlin.

[56] Deuse, J., Busch, F., 2012. Zeitwirtschaft in der Montage, in: Lotter, B., Wiendahl, H.-P. (Eds.), Montage in der industriellen Produktion. Springer, Berlin, Heidelberg.

[57] Dilling, H.-J., 1978. Methodisches Rationalisieren von Fertigungsprozessen am Beispiel Montagegerechter Produktgestaltung. Dissertation, Darmstadt.

[58] Dilling, H.-J., Rauschenbach, T., 1975. Rationalisierung und Automatisierung der Montage. VDI-Verlag, Düsseldorf.

[59] Eversheim, W., 1992. Organisation in der Produktionstechnik 4. Fertigung und Montage, 2nd ed. VDI-Verlag, Düsseldorf.

[60] Fischer, S., 1972. Mechanisierung der Montage in der Deutschen Demokratischen Republik. Fertigungstechnik und Betrieb 22 (9), 533–536.

[61] Korsakow, W.S., 1972. Typung von Montageprozessen und Automatisierung der Montagearbeit. Fertigungstechnik und Betrieb 22 (6), 384–387.

[62] Lotter, B., Schilling, W., 1994. Manuelle Montage: Planung - Rationalisierung - Wirtschaftlichkeit. VDI-Verlag, Düsseldorf.

[63] März, L., von Langsdorff, P., 2001. Flexibilität und Marktorientierung in der Montage, in: Westkämper, E., Bullinger, H.-J., Horváth, P., Zahn, E. (Eds.), Montageplanung - effizient und marktgerecht. Springer, Berlin, Heidelberg, pp. 3–10.

[64] Meyer, H., 1963. Theoretische Betrachtungen zur Mechanisierung und Automatisierung der Montage. Wissenschaftliche Zeitschrift der Technischen Hochschule Karl-Marx-Stadt 5 (2), 79–88.

[65] Miese, M., 1972. Ermittlung der Fehlerquellen bei der Montagedurchführung. Unveröffentlichter Untersuchungsbericht am Laboratorium für Werkzeugmaschinen und Betriebslehre der TH Aachen.

[66] Miese, M., 1973. Systematische Montageplanung in Unternehmen mit Einzel- und Kleinserienproduktion, Aachen.

[67] No Author, 1963. Assembly engineering methods. Mach. Shop Mag. 24 (4), 184–191.

[68] Peffekoven, K.H., 1982. Planung und Steuerung des Montageablaufs in Unternehmen der Einzel- und Serienfertigung. Dissertation, Aachen.

[69] Reinhart, G., Schneider, B., 1996. Montage, in: Kern, W., Schröder, H.-H., Weber, J. (Eds.), Handwörterbuch der Produktionswirtschaft, 2 ed. Schäffer-Poeschel, Stuttgart, pp. 1235–1247.

[70] Schimke, E.-F., Hoeschen, R.-D., 1977. Auswahl der geeigneten Organisationsform in der Montage. Technische Zeitschrift für praktische Metallbearbeitung 71 (4), 136–142.

[71] Schreiber, A., 1971. Probleme der technisch-ökonomischen Analyse von Montageprozessen. Messen-Steuern-Regeln 14 (8), 297–301.

[72] Seidel, U.A., 1998. Verfahren zur Generierung und Gestaltung von Montageablaufstrukturen komplexer Serienerzeugnisse. Springer, Berlin, Heidelberg.

[73] Stöferle, T., Dilling, H.-J., Tauschenbach, T., 1974. Rationalisierung und Automatisierung in der Montage. Werkstatt und Betrieb 107 (6), 327–336.

[74] Stöferle, T., Dilling, H.-J., Rauschenbach, T., 1975. Rationelle Montage - Herausforderungen an den Ingenieur. VDI-Z 111 (15/16), 715–719.

[75] VDI, 1978. VDI-Richtlinie 2815:1978-05: Begriffe für die Produktionsplanung und -steuerung. VDI-Verlag.

[76] VDMA, 1993. Statistisches Handbuch für den Maschinenbau, Frankfurt Main.

[77] Woithe, G., 1967. Kritische Analyse der Montageprozesse in einigen Betrieben der DDR. Fertigungstechnik und Betrieb 17 (7), 390–393.

Biography

Tobias Bein is a research associate and PhD student at the Institute for Engineering of Products and Systems at the Otto von Guericke University Magdeburg. His research interests include methods for factory planning, assembly system design and production planning for flexible and highly variable manufacturing systems.

Ulf Bergmann is a postdoctoral researcher at the Institute for Engineering of Products and Systems at the Otto von Guericke University Magdeburg. His research interests include production system planning with participatory approaches for SMEs and organizational development in manufacturing.

Oliver Antons received M.Sc. degrees in both Mathematics and Economics, as well as a doctoral degree in Economics from RWTH Aachen University. He worked as a research associate at RWTH Aachen from 2018 to 2022, and is currently a postdoctoral researcher and group leader for Data-driven Decision Making at Otto von Guericke University Magdeburg. For more information, including his full profile and bibliography, visit https://oliver.antons.eu.

Prof. Dr. Julia Arlinghaus is Chair of Logistics and Supply Chain Management and Director of the Institute for Production and Supply Chain Management at the University of St.Gallen. Her research focuses on new technologies and business models in logistics and manufacturing, methods of artificial intelligence, and new approaches to planning and control for digital and circular value creation systems.

8th Conference on Production Systems and Logistics

From Paper Production to Traction Motors: Transfer Patterns for the FOREST Digital Twin Framework

Rosario Othen[1], Nicolaus Klein[2], Florian Pohlmeyer[1], Christian Möbitz[1], Jürgen Fleischer[2], Thomas Gries[1]

[1]*Institut für Textiltechnik, RWTH Aachen University, Aachen, Germany*
[2]*wbk Insitute of Production Science, Karlsruhe Institute of Technology (KIT), Germany*

Abstract

The paper industry is highly energy intensive and exposed to volatile energy markets and strict decarbonisation targets. Digital twins can couple operational data with engineering and simulation models for energy and CO_2eq management, yet many solutions remain domain-specific. This paper summarises FOREST (Framework for Resource, Energy, and Sustainability Treatment) as a modular digital twin framework that integrates plant data, Asset Administration Shell (AAS)-based models, and executable simulations to provide consistent energy and material flow views, time- and product-resolved CO_2eq indicators, and scenario evaluations. Based on an industrial tissue implementation, three domain-independent architectural patterns are derived: flow and state tracking, asset-centric semantics, and co-execution of models and services. These patterns are conceptually mapped to electric traction motor manufacturing and End-of-Use (EoU) testing within industrial data spaces (IDS), illustrating how plant-level optimisation logic can be extended towards lifecycle-oriented decisions on value-retention strategies.

Keywords

Digital twin; CO_2eq accounting; Electric traction motor manufacturing; Paper production; Data management

1. Introduction

The decarbonisation of manufacturing is driven by two intertwined developments: the electrification of processes and products, and the increasing use of data-driven methods for monitoring, optimisation and reporting. Energy- and emission-intensive industries such as paper production must reduce fuel use, integrate volatile renewable electricity and provide verifiable product carbon footprints [1–4]. At the same time, electrified products such as traction motors are deployed at scale, creating new requirements for lifecycle assessment, EoU decisions and circular economy strategies [5–7].

Digital twins are key enablers of this transition because they couple operational data with engineering and simulation models for diagnostics, forecasting and scenario analysis [8]. In practice, many solutions remain limited to specific assets or platforms and focus on either production or downstream lifecycle services. A consistent, semantics-based chain from sensor data to lifecycle-oriented services is often missing. This limits comparability across plants and complicates integration into cross-organisational data spaces [9–12].

A Framework for Resource, Energy, and Sustainability Treatment (FOREST) is introduced as a modular digital twin framework. By combining automation data, semantic asset models and simulation units, FOREST generates energy and material flow views, time- and product-resolved CO_2eq indicators and

DOI: https://doi.org/10.15488/21004
ISSN: 2701-6277

scenario evaluations in tissue manufacturing, providing actionable guidance while satisfying transparency and auditability requirements. [10,13,14] In parallel, IDS and AAS initiatives are establishing federated infrastructures for sovereign data exchange along supply chains, which are particularly relevant for traction motors when deciding between remanufacturing, reuse and recycling at EoU [15,16,12,17].

Against this background, this paper asks how a framework such as FOREST can be generalised from its original process-industry context to discrete assembly and EoU services, and how it can be embedded into data-space-based infrastructures. The contributions of this work are threefold:

1. FOREST is summarised as a domain-agnostic digital twin architecture with core layers for data acquisition, semantic asset modelling, model integration and application services [10,13].
2. Experiences from an industrial tissue implementation are used to derive domain-independent patterns, including consistent flow and state tracking, asset-centric semantics and the co-execution of models and services [13,14].
3. These patterns are mapped to a traction motor use case in which an EoU test service operates on data provided via an IDS, illustrating how FOREST concepts can support circular strategies in discrete manufacturing [18,6,19,7].

To address this gap, this paper applies a pattern-based generalisation approach. The FOREST architecture, as implemented in an industrial tissue mill, is analysed to identify recurring solution elements that are not specific to fibre-based processing. These elements are generalised into three reusable patterns: flow and state tracking, asset-centric semantics, and co-execution of models and services. The patterns are then conceptually instantiated for a representative discrete-manufacturing use case, electric traction motor production with EoU testing. This demonstrates how the same architectural logic supports lifecycle-oriented services within IDS. The mapping specifies transferable design decisions and required interfaces rather than reporting a completed deployment.

The remainder of this paper is organised as follows. Section 2 reviews related work on digital twins for energy and CO_2eq management, industrial data ecosystems and AAS-based interoperability, and traction motor manufacturing and EoU services. Section 3 describes the FOREST framework and derives the three domain-independent patterns. Section 4 positions IDS as the federated environment for cross-organisational sharing of FOREST-generated information. Section 5 outlines the traction motor EoU test scenario as a pattern instantiation. Section 6 discusses limitations and practical deployment constraints, and Section 7 concludes with an outlook and concrete validation steps.

2. Background and Related Work

Digital twins, industrial data ecosystems and circularity-oriented services have been widely studied, but largely in separate communities. Process industry research focuses on model-based optimisation of continuous plants, including energy systems and CO_2eq management [1–4]. Manufacturing science addresses equipment, line and product twins in discrete assembly [5,6,19]. Parallel to this, IDS and AAS-based semantics have emerged from the Industrie 4.0 and Gaia-X context as a basis for sovereign, interoperable data exchange along supply chains [15,16,12,20,17]. In traction motor research, design and production aspects are increasingly linked to remanufacturing and recycling concepts, yet EoU testing is often treated as a stand-alone diagnostic step rather than as part of a continuous, data-driven lifecycle twin [19,7].

This paper connects these strands. Section 2.1 summarises the characteristics of fibre-based web production and reviews digital twin approaches for energy and CO_2eq management. Section 2.2 introduces industrial data ecosystems and the AAS as enablers of semantic interoperability and federated data sharing. Section 2.3 outlines traction motor production, information modelling and EoU testing, with a focus on how production and lifecycle data can inform R-strategies [21] such as remanufacturing, reuse and recycling.

2.1 Digital Twins for Energy and CO_2eq Management in Manufacturing

The paper industry is energy-intensive and characterised by tightly coupled, multi-stage processes. Thermal units such as hoods and cylinders dominate fuel consumption and direct CO_2eq, while drives, vacuum systems and auxiliaries add substantial electrical demand. Small changes in product mix, operating strategy or technology configuration can therefore significantly affect specific energy use and carbon footprint [1,2].

Although mills collect large amounts of process and energy data via automation systems, historians and business applications, information is often fragmented and primarily used for aggregated monthly reporting. Transient losses, cross-machine interactions and product-specific footprints remain difficult to analyse, and systematic scenario evaluation for electrification or fuel switching is rare due to missing operational models. [3]

State-of-the-art reviews highlight the largest levers for data and artificial intelligence enabled savings in mechanical refining, dewatering and drying, alongside condition monitoring and maintenance improvements [4]. Digital twins address these gaps by linking plant data with physics-based and data-driven models to create executable representations of selected assets or process sections [8]. Existing work in the paper industry typically focuses on unit operations or quality optimisation and is often tied to proprietary platforms, which limits reuse and cross-vendor integration [10,11,3].

With FOREST, energy, resource, and emission flows are represented through modular, interoperable digital-twin components that can be instantiated and combined to reflect site-specific mill configurations [10]. A pilot in tissue manufacturing showed that FOREST can realise tangible energy savings and support energy flexibility [14,22]. The underlying requirements, i.e. traceable energy and material flows, asset-centric semantics, integrated executable models and transparent CO_2eq calculations [13], are not specific to paper and apply to discrete domains such as traction motor production. This motivates the generalisation of FOREST beyond paper industry.

2.2 Industrial Data Ecosystems and Semantic Interoperability

Industrial data ecosystems constitute socio-technical environments in which multiple actors generate, exchange, and use data across organizational boundaries. Their central feature is a federated architecture that preserves data sovereignty while enabling secure and interoperable data sharing. This is achieved through standardized interfaces, semantic models, and shared governance mechanisms [12]. Core building blocks include data sources, integration connectors, identity and access management, semantic interoperability components, and data governance services. Together these components ensure that heterogeneous information from different systems can be exchanged and interpreted reliably [15].

A recurring challenge in industrial practice is the heterogeneity, inconsistency, and limited availability of data, which often inhibit cross-organizational collaboration. To address these limitations, current industrial initiatives emphasize standardization [16,23,9]. The AAS has emerged as a central interoperability concept in industrial data ecosystems. It is often described as an integration connector because it provides a unified semantic representation of assets across their lifecycle [16]. According to its formal definition, the AAS enables secure access to asset-related information through structured submodels, standardized identifiers, and machine-readable semantics [24]. These characteristics make the AAS a key enabler for consistent data exchange and reduce manual mapping efforts across heterogeneous system landscapes.

Data ecosystems have become essential for industrial transformation, particularly in sustainability-related applications where traceability, transparency, and trustworthy data provision are required across supply chains. They allow the integration of production, quality, and environmental data, enabling advanced analyses and evidence-based decision-making. However, the successful implementation still depends on scalable, partially automated data connectors and harmonized semantic standards, since many enterprise systems are not natively compatible with interoperable architectures. [25–28]

2.3 Electric Traction Motor Manufacturing and Lifecycle

Electric traction motors are a key technology for battery electric vehicles and strongly shape performance, efficiency and sustainability. Permanent magnet synchronous motors, especially interior permanent magnet designs with hairpin windings, dominate automotive traction due to high power density, high-speed efficiency, compact design and low noise. Their laminated stator, magnet rotor and integrated housing create high product and process complexity, requiring detailed process-chain analysis across rotor and stator production and final assembly. [5,6]

Original equipment manufacturers demand end-to-end documentation and traceability for safety- and quality-critical components. This requires continuous data capture across suppliers, manufacturing, inspection, and logistics. To achieve motor-level traceability aligned with ISO 9001 and ISO 26262, legacy IT systems must be integrated into a consistent production landscape [19,29]. This makes information modelling and service-oriented IT integration central to traction motor production. In this context the AAS provides a standardised digital twin implementation. It utilizes modular submodels for technical, production, traceability, and bill-of-materials data. Industrial pilots indicate that AAS-based approaches support scalable, reconfigurable production and consistent real-time mapping of product, process, and machine parameters [29].

Beyond production, lifecycle-oriented information modelling remains essential. It enables a unified and traceable representation of product architecture, functional behaviour, manufacturing and test processes, and lifecycle data. The remanufacturing of electric powertrains relies on integrating condition information from the use phase, such as thermal loads, vibration exposure, operating cycles, and fault histories, because key components including magnets, hairpin windings, insulation systems, and bearings are difficult to recover and highly sensitive to prior loading and degradation mechanisms [7].

The EoU phase is central to techno-economic and environmental assessment. EoU testing, including insulation resistance, torque–speed evaluation, vibration-based bearing diagnostics and thermal analysis, provides the empirical basis for assessing technical condition, determining remanufacturing potential and feeding back into design, quality assurance and lifecycle optimisation. Sustainable traction motor systems therefore require systematic integration of production, operation, EoU testing and remanufacturing within a consistent, AAS-based, traceable information model. [7]

3. The FOREST Digital Twin Framework

FOREST (Framework for Resource, Energy, and Sustainability Treatment) is conceived as a domain-agnostic, yet industry-proven, digital twin architecture for energy and CO_2eq management. It provides a structured way to connect heterogeneous data sources, semantic asset models and executable simulations to generate actionable indicators and decision support services. While it has been developed and validated initially in tissue production, its concepts are deliberately formulated in a way that allows transfer to other manufacturing domains [10,11,13].

3.1 Objectives and Scope

The primary objective of FOREST is to enable real-time and historical analysis of energy and CO_2eq performance at asset, line and product level, and to support scenario-based evaluation of technology and operating changes. Concretely, the framework aims to

- capture and harmonise production, energy and quality data from existing automation and IT systems,
- map these data onto asset-centric semantic models that represent the structure and behaviour of the production system,
- couple the semantic models with executable simulation units to explore alternative configurations and operating windows, and

- derive key performance indicators for energy use and CO_2eq emissions, suitable for operational guidance and external reporting.

In doing so, FOREST explicitly targets use cases such as root-cause analysis of energy deviations, identification of best-practice operating windows, evaluation of electrification, and assessment of product-specific footprints along different routing options or life-cycle scenarios.

3.2 Architectural Layers and Data Management

FOREST is conceived as a gate-to-gate, mill- or line-level advisory framework. It links process data, engineering models and energy and CO_2eq accounting, while remaining deployable in existing brownfield environments, modular, evolvable and economically viable. Monolithic one-off solutions and purely cloud-based system landscapes are therefore excluded. Instead, a layered microservice architecture is adopted in which FOREST is decomposed into a small set of coarse-grained services that can be deployed, scaled and updated independently.

The services follow the responsibility clusters data acquisition and integration, semantic modelling and asset management, simulation and scenario analysis, energy and CO_2eq calculation, and visualisation and decision support, complemented by cross-cutting services and tools. Each service owns its data and exposes a stable interface. The services are arranged in layers that separate core data and semantics, analytical services and user-facing decision support, as depicted in Figure 1, in line with the entity-based digital twin reference model for manufacturing [30]. An AAS-based semantic layer and Functional Mock-up Unit (FMU)-based model integration are employed so that assets, data sources and simulation models can be linked consistently across services [31]. The layers can be distinguished as follows:

1. **Data acquisition & semantic modelling and asset management**: The need to connect control systems and production IT, and to cope with high-frequency signals and inconsistent tag naming, calls for a data integration layer with connectors, buffering, and basic data quality services. Requirements for a consistent representation of machines, sections, sensors and energy systems result in AAS based asset models and an asset registry that provide a shared semantic layer and enable reuse of digital twin descriptions.
2. **Simulation and scenario analysis**: Requirements for offline and online studies of speed changes, retrofits and energy supply options lead to a simulation engine that loads and executes FMUs, manages parameter sets and orchestrates multi model scenarios.
3. **Energy and CO_2 calculation**: Requirements to convert energy flows into CO_2eq indicators under different system boundaries lead to a dedicated energy and emissions calculation service.
4. **Visualisation and decision support**: Requirements for dashboards, key performance indicators, operator guidance and reporting lead to a visualisation layer with interfaces for web dashboards, reports and external applications.
5. **Services and tools**: Requirements for configuration, exploration and integration lead to supporting components such as explorers, data services, optimisation tools and message brokers that interact with all layers and extend the FOREST core for specific tasks.

Supporting services and tools implement data management and governance, configuration management, user management and technical infrastructure functions. The figure also shows how the FOREST services interact with existing shopfloor and enterprise systems as well as external energy-market data sources via secure data connections, thereby preparing the architecture for integration into IDS.

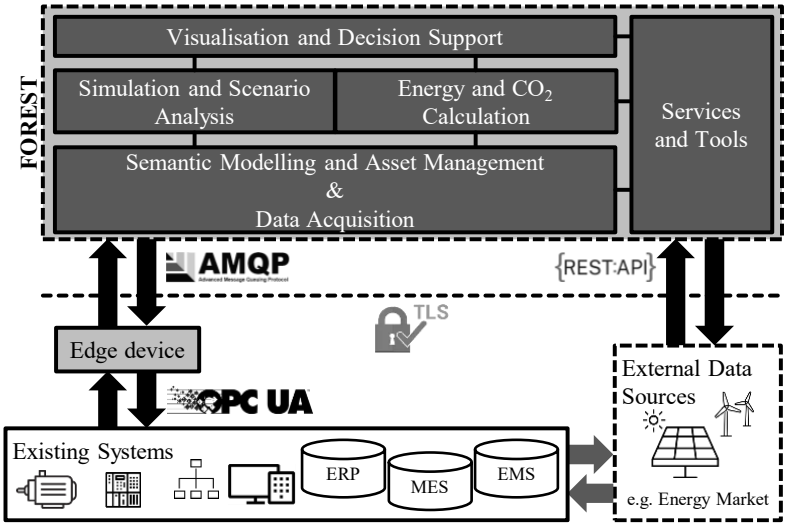


Figure 1: FOREST architecture with layered services

3.3 Domain-Independent Patterns

A set of domain-independent patterns can be abstracted from Figure 1 that enable the transfer of FOREST from paper manufacturing to traction motor manufacturing and data-space-based EoU services:

Flow and state tracking across assets and time, in which material, energy and footprint flows are represented consistently along the process topology and associated with operating states and events. In continuous paper production this corresponds to mass and energy balances across sections, whereas in traction motor assembly it corresponds to component- and process-step tracking along alternative routing variants and test stages.

Asset-centric semantics, realised by AAS-based asset models and the asset registry. This establishes a shared, machine-readable language for internal and cross-organisational applications. It can be used when a traction motor producer exposes product and process twins into a federated data space or when an EoU test service consumes lifecycle data to derive remanufacturing or recycling recommendations [16,17].

Co-execution of models and decision-support services. Data streams from the acquisition layer are coupled with FMU-based and data-driven models as well as with calculation services via stable interfaces. In paper production, this allows energy demand and CO_2eq effects of alternative drying concepts or energy supply options to be estimated. In traction motor production, analogous configurations can estimate residual lifetime, remanufacturing effort or CO_2eq impacts of different R-strategies [21] based on combined production, usage and EoU test data.

Taken together, these patterns show that the FOREST architecture is not limited to a specific process type but can act as a reusable digital-twin backbone that can be instantiated across continuous and discrete domains and prepared for secure integration into IDS.

4. Data Spaces

Building on the architectural patterns described above, this section positions IDS as the federated environment in which the FOREST framework can extend its functionality beyond plant-level optimisation. While FOREST harmonises operational, energy and CO_2eq data through semantic asset models, data spaces provide the decentralised infrastructure required to make these information assets securely available to downstream lifecycle services.

Current manufacturing IT landscapes are characterised by fragmented systems, bilateral interfaces and limited interoperability, which hinder the integration of operational and environmental data across

organisational boundaries [32–34]. Data spaces address these limitations through decentralised architectures, shared governance and interoperable standards that enable sovereign data exchange across value chains [35,36,17]. In this context, the AAS functions as a semantic interoperability layer, allowing FOREST-generated indicators, events and models to be exposed in a machine-readable and standardised form [15,20].

Established reference architectures, such as the IDS RAM with its security, certification and governance perspectives [12] and the Data Spaces Support Centre (DSSC) building blocks for identity, metadata and connector services, provide the structural basis for integrating FOREST into federated data ecosystems (see Figure 2). Through these components, production data and derived indicators can be shared with lifecycle applications without compromising data ownership or confidentiality.

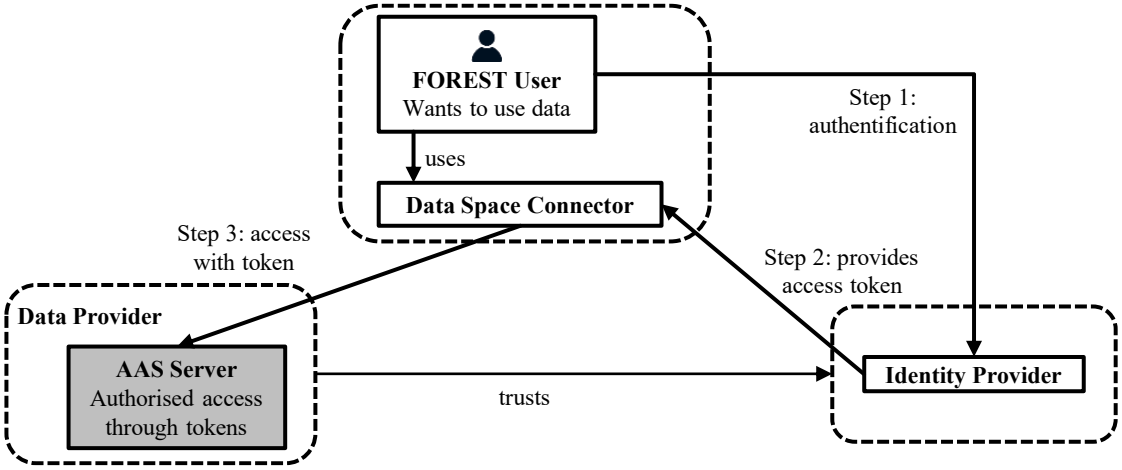


Figure 2: Conceptual integration of FOREST and AAS-based twins into an IDS

For the traction motor EoU test, data spaces enable the controlled combination of production, quality and material data with lifecycle-relevant information to derive robust R-strategies. Requirements regarding data quality, semantic consistency and reproducible calculation rules ensure that CO_2eq and performance indicators remain trustworthy when transferred between actors [18].

In this role, data spaces operationalise the cross-domain transfer envisioned in this paper: they create the conditions under which a generalised digital twin framework like FOREST can support both operational guidance in manufacturing and lifecycle-oriented decision-making in traction motor manufacturing. In this context, FOREST can be seen as a service of the data space platform.

5. Conceptual EoU Testing in Traction Motor Production Using FOREST Patterns

In the traction motor context, FOREST is applied through a pattern instantiation rather than by transferring a fibre-specific solution. The goal is to consistently link product identity, process and quality states, and energy/CO_2eq indicators across multi-stage manufacturing chains, for example for permanent magnet synchronous machines manufactured with hairpin technology. The three patterns structure this mapping. Flow and state tracking relates material and energy flows to process events and quality states across stations and batches. Asset-centric semantics represent motors, components, and test artefacts as interoperable, AAS-based digital twins with shared identifiers and controlled vocabularies. Co-execution couples these semantic representations with simulation and analytics services to compute derived indicators and support scenario-based assessments.

EoU testing is treated as a lifecycle event that generates decision-relevant evidence for value retention. Typical test sequences include insulation resistance measurements, torque-speed characterisation, vibration-based bearing diagnostics, and thermal measurements. These signals become decision-relevant only when contextualised with manufacturing and usage history. Relevant context includes component genealogy,

process deviations, quality gates, repairs, and operating conditions. In the proposed architecture, this linkage is realised through AAS-described data structures and governed data exchange via IDS. This enables role-specific and cross-organisational access without centralising data.

Decision-support services are parameterised by EoU findings and contextual histories to evaluate value-retention pathways such as remanufacturing, component reuse, and recycling. The same service layer can estimate associated effort, expected residual lifetime, and feasibility constraints. It can also quantify comparative CO_2eq impacts. The focus therefore extends from production-only optimisation to lifecycle-integrated assessment, while retaining the same architectural interfaces and pattern logic.

To illustrate the quantitative impact of the proposed approach, a representative traction motor scenario is considered. Life-cycle assessment studies on electric traction motors and remanufactured electrical machines show that manufacturing-related impacts can amount 300–500 kgCO_2eq [37], while remanufacturing can substantially reduce these emissions; for example, one study reports emission reductions of 26.75% to 65.13% depending on the remanufacturing case [38]. This implies that, depending on the dedicated assessment conducted during the end-of-use testing phase, a substantial potential for CO_2 savings can be achieved. In particular, significant reductions are possible when complete components, such as the magnet-equipped rotor, which is associated with a high carbon footprint of approximately 250 kgCO_2eq [39], can be reused following the replacement of wear-prone elements such as bearings.

6. Discussion

The results from the tissue pilot and the traction motor EoU service illustrate that the digital twin framework originally developed for fibre-based web production can be generalised to a discrete manufacturing domain and integrated into data-space-based infrastructures. The three patterns from the tissue implementation, namely flow and state tracking, asset-centric semantics and co-execution of models and services, are transferable to traction motor production and EoU services. However, domain-specific specialisations such as magnet composition, insulation systems or dedicated test submodels remain necessary.

Embedding FOREST into IDS highlights the central role of standardisation and governance. AAS-based semantics and IDS connectors provide a solid interoperability stack yet deploying them in brownfield environments remains effort-intensive due to modelling, connector setup, and data governance, which can be especially challenging for SMEs. The EoU service concept illustrates how plant-level twins can support lifecycle decisions such as R-strategy selection and product carbon footprint communication, but practical rollouts must address incomplete data and organisational barriers to cross-company sharing. Hybrid approaches that combine model-based estimation, partial measurements, and expert knowledge will therefore often be necessary. Overall, FOREST can connect shop-floor optimisation with lifecycle services, provided interoperability, semantics, and governance are developed and maintained.

7. Conclusion and Outlook

The FOREST digital twin framework is conceptually adapted from fibre-based web production to discrete manufacturing and lifecycle services. It is achieved via three reusable patterns, namely flow and state tracking, AAS-based asset-centric semantics, and co-execution of models and services. The framework is instantiated for traction motor manufacturing and EoU testing in an IDS context to link operational indicators with lifecycle diagnostics for value-retention decisions and CO_2eq assessment.

Key remaining challenges concern the effort required for AAS-based semantic modelling, IDS-compliant connectors, and governance in heterogeneous brownfield environments. Future work will focus on industrialisation and scaling. This includes traction-motor-specific AAS submodels, integration with MES and quality and test infrastructures, and validation using representative motor populations. In parallel, hybrid

physics- and data-driven methods under incomplete data will be investigated, together with additional domain applications and alignment with product carbon footprint and digital product passport regulations.

Future work should operationalise this instantiation through implementation of traction-motor-specific AAS submodels (product, process, test, and condition-history). This includes integrating existing MES/quality/test infrastructure and connector-based provision within an IDS-compliant data space. In addition, calculation rules for lifecycle- and remanufacturing-related indicators are defined and validated under conditions of incomplete data. The approach is evaluated on representative motor populations. The evaluation considers integration effort, runtime performance, and decision-quality impacts compared with baseline EoU testing workflows.

Acknowledgements
This work is partially supported by German Federal Ministry for Economic Affairs and Energy of Germany (BMWE) in the scope of the FOREST project (03EN2095 A-G) and the Mittelstand-Digital Zentrum Smarte Kreisläufe (01MF23001B) project.

References

[1] Holik, H. (Ed.), 2013. Handbook of paper and board, 2.th ed. Wiley-VCH, Weinheim, Germany.

[2] Naujock, H.-J., Blechschmidt, J. (Eds.), 2021. Taschenbuch der Papiertechnik, 3rd ed. Hanser, München.

[3] Othen, R., Padberg, J., Möbitz, C., Gries, T., 2025. Digital Transformation in the Paper Industry: Assessing Maturity, Challenges, and Opportunities. Sustainability 17 (2), 770.

[4] Viitala, R., Miettinen, M., Marquez, R., Hämäläinen, A., Karhinen, A., Barrios, N., Gonzalez, R., Pal, L., Jameel, H., Holmberg, K., 2025. Integration of artificial intelligence and sustainable energy management in the pulp and paper industry: A path to decarbonization. Renewable and Sustainable Energy Reviews 218, 115809.

[5] Cai, W., Wu, X., Zhou, M., Liang, Y., Wang, Y., 2021. Review and Development of Electric Motor Systems and Electric Powertrains for New Energy Vehicles. Automot. Innov. 4 (1), 3–22.

[6] Fleischer, J., Ceglarek, D., Franke, J., Herrmann, C., 2025. Production technologies and systems for electric mobility. CIRP Annals 74 (2), 1047–1072.

[7] Scholz, J., Klein, N., Kößler, F., Fleischer, J., 2025. Systematic Digital Twin-Based Development Approach for Holistic Sustainable Electric Traction Motors. Sustainability 17 (6), 2518.

[8] Pal, P.K., Hens, A., Behera, N., Lahiri, S.K., 2025. Digital twins: Transforming the chemical process industry: A review. Can J Chem Eng 103 (8), 3611–3636.

[9] Jacoby, M., Volz, F., Weißenbacher, C., Stojanovic, L., Usländer, T., 2021. An approach for Industrie 4.0-compliant and data-sovereign Digital Twins. at - Automatisierungstechnik 69 (12), 1051–1061.

[10] Juhlin, P., Othen, R., Zehnpfund, A., Song, C., Schlake, J., Schmeiser, A., Kratzer, K., Sejdija, J., Kayser, P., 2024. Open Reference Architecture for Sustainable Papermaking based on Industrial Interoperability Standards and Cloud-Native Technologies, in: 2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (EFTA). Conference on Emerging Technologies and Factory Automation, Padova, Italy. 10 - 13 September 2024. IEEE, Piscataway, NJ.

[11] Juhlin, P., Schlake, J., Song, C., Zehnpfund, A., Grüner, S., Schmeiser, A., Kratzer, K., Othen, R., Sejdija, J., Kayser, P., 2025. CRADLE: Cross-Vendor Asset Digital Twin Architecture using Industrial Interoperability Standards with Application to Paper Manufacturing, in: IEEE ETFA 2025. 2025 IEEE 30th International Conference on Emerging Technologies and Factory Automation (ETFA). 2025 IEEE 30th International Conference on Emerging Technologies and Factory Automation (ETFA), Porto, Portugal. 09 - 12. September 2025. IEEE, Piscataway, NJ.

[12] Otto, B., Steinbuß, S., Teuscher, A., Bader, 2022. IDS Reference Architecture Model, Berlin. https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4/front-matter/frontmatter.

[13] Othen, R., Pohlmeyer, F., Zehnpfund, A., Fidêncio, A.X., Song, C., Schlake, J., Schmeiser, A., Kayser, P., Möbitz, C., Gries, T., 2025. Entwicklung einer Referenzarchitektur für das Monitoring und Optimieren von Energie und CO2eq-Emissionen in der Papierindustrie, in: Automation. 26. Leitkongress der Mess- und Automatisierungstechnik Automation 2025 - Human-centric Automation, Baden-Baden. 01.-02.07.2025. VDI Wissensforum GmbH, Düsseldorf, pp. 683–710.

[14] Othen, R., Sejdija, J., Prinz, M., Zehnpfund, A., Lauricella, M., Kayser, P., Gries, T., 2025. Decarbonisation in the Paper Industry: A Scalable Digital Approach for CO_2eq Transparency, Energy Efficiency, and Flexibilisation. atp magazin - Magazin für Automatisierung und digitale Transformation 67 (8), 40–51.

[15] Data Spaces Support Centre (DSSC), Fraunhofer-Institut für Software und Systemtechnik ISST, 2025. Data Spaces Blueprint v2.0 - Home. Data Spaces Support Centre (DSSC); Fraunhofer-Institut für Software und Systemtechnik ISST.

[16] 2023. Deutsche Normungsroadmap Industrie 4.0: Version 5, Berlin, Offenbach am Main.

[17] Reiberg, A., Niebel, C., Kraemer, P., 2022. Was ist ein Datenraum?: Definition des Konzeptes Datenraum. acatech - Deutsche Akademie der Technikwissenschaften, München.

[18] Europäische Kommission, 2021. EMPFEHLUNG (EU) 2021/2279 DER KOMMISSION vom 15. Dezember 2021 zur Anwendung der Methoden für die Berechnung des Umweltfußabdrucks zur Messung und Offenlegung der Umweltleistung von Produkten und Organisationen entlang ihres Lebenswegs.

[19] Klein, N., Herm, M., Matussek, R., Fraider, F., Kößler, F., Fleischer, J., 2025. Informationsmodelle und Integrationskonzepte für die Produktion elektrischer Traktionsmotoren. Zeitschrift für wirtschaftlichen Fabrikbetrieb 120 (5), 314–319.

[20] Plattform Industrie 4.0, 2019. Digitale Ökosysteme global gestalten. Leitbild 2030 für Industrie 4.0, Berlin.

[21] Cramer, J., 2017. The Raw Materials Transition in the Amsterdam Metropolitan Area: Added Value for the Economy, Well-Being, and the Environment. Environment: Science and Policy for Sustainable Development 59 (3), 14–21.

[22] Sejdija, J., Schemm, R., Lauricella, M., Maurer, F., Prinz, M., Dickler, S., Kuperjans, I., 2025. Regulation-aware optimization for bi-valent industrial processes. Energy Inform.

[23] 2023. Deutsche Normungsroadmap Circular Economy, Berlin.

[24] IEC International Electrotechnical Commission, 2023. Asset Administration Shell for industrial applications - Part 1: Asset Administration Shell structure.

[25] Canepa, A., Fite Galan, C., Sabbata, P. de, Ciaccio, G., Desabbata, P., Fantin, V., Cortesi, S., 2024. Data collection for DPP - Policy Brief Deliverable D5.6 - Performance evaluation report of TC pilot.

[26] IW Consult GmbH; Fraunhofer-Institut für System- und Innovationsforschung ISI, 2024. Zukünftige Unterstützungsbedarfe des Mittelstandes in der digitalen Transformation: Studie für den DLR Projektträger im Rahmen der Projektträgerschaft "Mittelstand-Digital", Köln.

[27] Gries, C.-I., Tenbrock, S., 2023. Internet of Things (IoT): Vernetzte Geräte und Maschinen im Mittelstand, Bad Honnef.

[28] Jensen-Abieva, E., Roedding, T.L., Mühlenen, A. von, Alcayaga, A., 2025. The Economics of Digital Product Passports From Materials to Market. CircularTech - the Global Digital Product Passport Community, Köln.

[29] Klein, N., Wirth, F., Fleischer, J., 2023. Methodology for the Implementation of a Consistent Information Model for the Electric Drives Production, in: 2023 13th International Electric Drives Production Conference (EDPC). 2023 13th International Electric Drives Production Conference (EDPC), Regensburg, Germany. 29.11.2023 - 30.11.2023. IEEE, pp. 1–8.

[30] ISO, 2021. Automation systems and integration - Digital twin framework for manufacturing: Part 2: Reference architecture. International Organization for Standardization, Geneva.

[31] Modelica Association Project FMI (Ed.), 2023. Functional Mock-up Interface Specification: Version 3.0.1. Modelica Association Project FMI. https://fmi-standard.org/docs/3.0.1/. Accessed 16 November 2023.

[32] Bonfiglio, F., 2021. Vision & Strategy. GAIA-X European Association for Data and Cloud, Brüssel.

[33] Dumss, S., Weber, M., Schwaiger, C., Sulz, C., Rosenberger, P., Bleicher, F., Grafinger, M., Weigold, M., 2021. EuProGigant – A Concept Towards an Industrial System Architecture for Data-Driven Production Systems. Procedia CIRP 104, 324–329.

[34] Zeller, V., Schuh, G., Stich, V., 2022. Einleitung - Digitalisierungs- und Informationsmanagement für produzierende Unternehmen, in: Schuh, G., Zeller, V., Stich, V. (Eds.), Digitalisierungs- und Informationsmanagement. Handbuch Produktion und Management 9, vol. 1. Springer Vieweg, Berlin, Heidelberg, pp. 1–16.

[35] Garrels, K., Jochem, M., Jänicke, L., 2023. The AAS Dataspace for Everybody: An Architecture Example for a Simple Dataspace – purely based on the Asset Administration Shell Concepts, Berlin.

[36] Otto, B., Seidelmann, J., Schmelting, J., Sauer, O., 2023. Vorstudie Datenraum Manufacturing-X: Architektur, Basisdienste und Organisation unter Berücksichtigung der Spezifika der ausrüstenden Industrie.

[37] Nordelöf, A., Tillman, A.-M., 2018. A scalable life cycle inventory of an electrical automotive traction machine—Part II: manufacturing processes. The International Journal of Life Cycle Assessment 23 (2), 295–313.

[38] Liu, M., Chen, L., Sheng, X., Yang, Y., Yu, F., Li, Y., Yuan, X., Li, Y., Wang, Q., Ma, Q., 2023. Dynamic simulation of life cycle environmental benefits of remanufacturing asynchronous motors to permanent magnet synchronous motors. Journal of Cleaner Production 426, 138932.

[39] Klein, N., Franken, L., Heim, M., Kößler, F., Dönges, B., Fleischer, J., 2025. Assessment of Product Carbon Footprint Reduction Potential Using Lightweight Rotor Components for Electric Traction Motors, in: Kohl, H., Seliger, G., Dietrich, F., Vien, H.T. (Eds.), Decarbonizing Value Chains. Springer Nature Switzerland, Cham, pp. 91–98.

Biography

Rosario Othen (*1996) has been a Research Associate at the Institut für Textiltechnik (ITA) of RWTH Aachen University since 2021. Since 2022, he has served as a member of the advisory board of ZELLCHEMING e.V. and as Chair of its Technical Committee DATA. His expertise lies in sustainability management through digital twins for energy and carbon footprint optimization, paper technology, and empowering small and medium-sized enterprises (SMEs) to implement digital transformation strategies.

Nicolaus Klein (*1997) has been a Research Associate at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT) since 2022. His main areas of expertise include process development and modelling for sustainable electric motor production, as well as digital twin concepts and information models for resource-efficient manufacturing of permanent magnet rotors and electric traction motors.

Florian Pohlmeyer (*1998) has been a Research Associate at the Institut für Textiltechnik (ITA) of RWTH Aachen University. Additionally, he has been the Head of Digitalisation at the ITA since 2023. His main areas of expertise include data ecosystems and sustainability in the fiber-based process industries with a special focus on transformation processes in small and medium-sized enterprises

Jürgen Fleischer (*1961) obtained his doctorate at the Institute of Production Science (wbk) in 1989. From 1992 on, he held several leading positions in industry before being appointed professor and head of the wbk at today's Karlsruhe Institute of Technology (KIT) in 2003. His current scientific research focuses on

intelligent production machines and components, the automation of immature processes and agile production systems.

Thomas Gries (*1964) is Professor at the Institut für Textiltechnik (ITA) of RWTH Aachen University and the Director of the ITA since 2001. He has authored and co-authored more than 850 scientific articles and held more than 160 presentations, mainly at international congresses.

8th Conference on Production Systems and Logistics

Global Warming Potential Assessment for Sustainable Procurement in Production Supply Chains

Kunal Chaudhari[1], Oussama Etlijani[1], Clemens Gonnermann[1], Aljoscha Hieronymus[1], Maximilian Holland[1], Sönke Maatsch[2], Anja Munzke[1], Frida Schulze[1]

[1]*Fraunhofer Institute for Casting, Composite and Processing Technology IGCV, Augsburg, Germany*
[2]*Institute of Shipping Economics and Logistics ISL, Bremen, Germany*

Abstract

As climate regulations tighten, investor expectations rise, and customers demand transparency, the ability to measure and understand supply chain emissions has become essential. Beyond enabling targeted decarbonization, emissions data now plays a critical role in assessing a company's long-term resilience and competitiveness. The Global Warming Potential (GWP) assessment approach presented in this paper provides a standardized, transparent approach for measuring greenhouse gas emissions in terms of carbon dioxide equivalents along supply chains, enabling consistent comparisons of production options and procurement strategies. The approach is based on open-source Life Cycle Inventory (LCI) data and combines process- and country-specific emission factors to estimate GWP impacts for individual production and transport steps. Features of the approach can be used for raw material extraction, intermediate processing, final assembly, and transportation, identifying hotspots and prioritizing measures. Scenarios can be analyzed, making it possible to assess how alternative supplier locations, transport routes, or energy procurement decisions affect overall GWP and associated operational risks. By providing reproducible metrics and scenario comparisons, the approach supports evidence-based mitigation measures, more resilient procurement decisions, and sustainable procurement planning.

Keywords

GWP Assessment; Life Cycle Assessment; Industry; Manufacturing; Transportation

1. Introduction and motivation

The accelerating impacts of climate change require immediate, systemic interventions to meet the international climate objectives defined by the Paris Agreement. These goals commit nations to holding the increase in the global average temperature to well below 2 °C and pursuing efforts to limit it to 1.5 °C above pre-industrial levels [1]. This mandate has triggered a rapid, comprehensive economic transition, placing significant regulatory and market pressure on the highest-emitting sectors, specifically the power sector, industry, and transportation, which together account for 85,9 % of global CO_2 emissions, based on 2024 statistics [2]. As a result, companies are increasingly required to monitor and manage their environmental impact, often by adopting established assessment approaches and reporting standards to ensure transparency and alignment with global climate objectives [3].

Despite this growing requirement, organizations continue to face substantial challenges in delivering comprehensive climate reporting. While many companies can quantify emissions from their own operations

DOI: https://doi.org/10.15488/21005
ISSN: 2701-6277

(e.g., facility combustion and transportation) and from purchased energy (e.g., electricity or heat), their disclosures often remain confined to these direct sources [4, 5]. This limitation is particularly pronounced for small and medium-sized enterprises (SMEs), which may struggle to collect accurate data for their upstream and downstream activities, commonly referred to as value chain emissions [6]. Reliable and specific supply chain data are often difficult to obtain, and the lack of standardized, reproducible, and transparent approaches further complicates assessments [6, 7]. Current approaches often rely on fragmented, non-public data, leading to inconsistent, incomparable emission results that hinder robust decision-making. These approaches are implemented in software-based assessment tools, but their dependence on closed data sources and the lack of free tools often limit the ability to perform transparent and efficient GWP analysis.

To address this gap, this paper presents an approach for GWP assessment that uses open-source inventory data and country-specific factors to consistently quantify greenhouse gas (GHG) emissions across all stages of the supply chain. The approach supports comprehensive scenario analysis, enabling organizations to evaluate how strategic adjustments, such as relocating suppliers or modifying energy sources, alter their total GHG emissions, and serves as the methodological foundation for the development of a software-based assessment tool.

2. State of the art

2.1 Standards for environmental assessment

Existing guidelines already enable the standardized assessment of the GWP of companies and products. As a fundamental standard, ISO 14040/14044 specifies the conducting of a Life Cycle Assessment (LCA) [8, 9]. Figure 1 illustrates the procedure for the assessment of production sites in the phases LCI and Life Cycle Impact Assessment (LCIA).

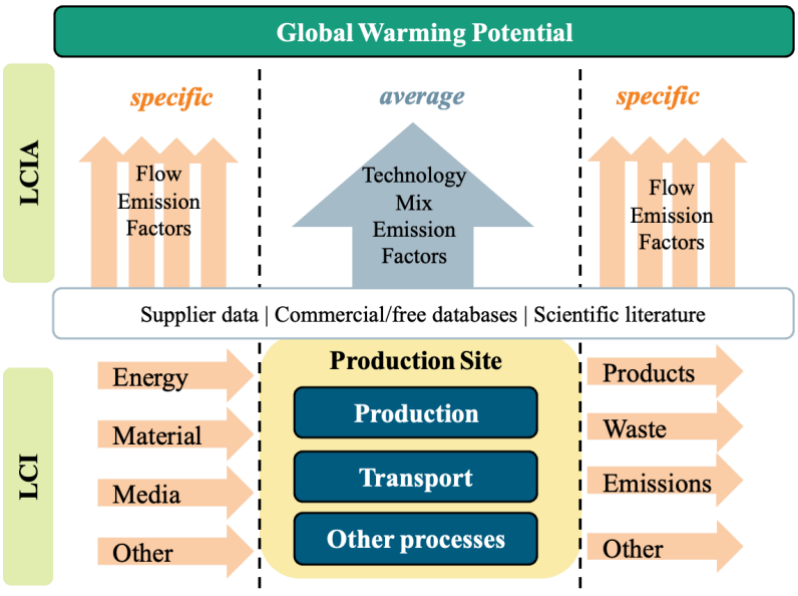


Figure 1: Overview of LCI/LCIA of production sites

In the LCI phase for all production, transportation, and other processes (e.g., storage), the required input and output flows, such as energy consumption, material use, emissions, and waste, are quantified and documented. In the subsequent LCIA phase, the environmental impacts, particularly the GWP, associated with these flows are evaluated using flow-specific emission factors, gathered from suppliers (primary data), commercial or open-source available LCA databases, and scientific literature (secondary data). Alternatively, if the flows cannot be quantified, for each process, average “technology mix” emission factors

can be used directly, leading to less specific results. The LCA standard defines the basic methodological approach and extensive documentation requirements. Nevertheless, the standard approach remains sufficiently accommodating to allow for the establishment of system boundaries or the selection of relevant processes and emission factors [8].

Standards compliant to ISO 14040/14044 define these methodological decisions more narrowly. In Europe, industry-specific standards in accordance with EN 15804 and Product Environmental Footprint can usually be used to calculate product-specific environmental impacts [10]. Focusing on GWP, ISO 14067/14064 can be used to calculate specifically the product carbon footprint or the corporate carbon footprint [11, 12]. Companies rely on experts to calculate GWP in accordance with one of the listed standards, which is a time- and cost-consuming process, with the need for holistic data gathering. This poses a particular challenge for SMEs. For this reason, various stakeholders (including research groups and industry associations) have developed tools designed to help companies quickly estimate the GWP of their products and their business.

2.2 Analysis of existing tools: GWP assessment in production

For the assessment of the GWP of product production, a wide range of calculation tools are available. They can be categorized regarding their accessibility, type, and sector specification (see Figure 2).

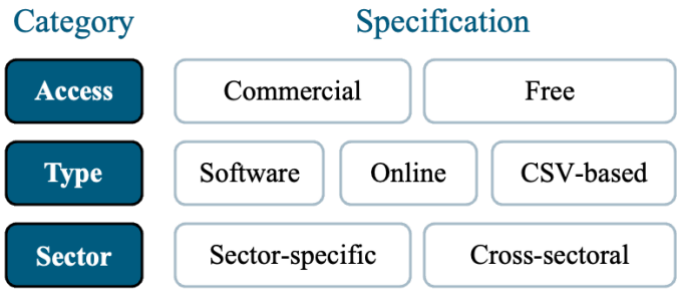


Figure 2: Categorization of GWP assessment tools

Tools include both commercial and freely accessible options, which may be designed for sector-specific applications or for use across multiple sectors. Commercial solutions typically exist either as dedicated software packages or as online platforms, whereas free tools are most often provided as web-based applications or CSV-based calculation sheets. Together, these options offer varying levels of detail, usability, and sector specificity for conducting GWP assessments.

Commercial tools allow the calculation of the GWP as well as further environmental impacts (e. g. acidification). They combine the use of commercial LCA databases with specific interfaces, adapted to the required level of detail per offered solution, ranging from detailed LCA to simpler GWP assessment. LCA databases provide extensive technology mixes and flow-specific data, enabling both sector-specific and cross-sectoral assessments. The adapted level of detail reduces the complexity for simpler tools, although they remain expert tools. However, industry-leading tools with extensive databases and functionalities are costly, while more affordable tools provide limited data and functionality. This restricts access to commercial tools, particularly for SMEs [13].

The wide range of *free tools* has already been analyzed in various studies, identifying their different characteristics [14–16]. *Cross-sectoral tools* are primarily aligned with ISO 14067/14064 and enable detailed assessments (e.g., “GHG Emissions Calculator”) [17]. However, they require extensive LCI data from users and generally do not offer technology mix data to support data acquisition. This also implies that unknown process data will result in gaps in the assessment, thereby limiting the requirement for complete data availability.

Sector-specific tools often provide technology mix data, and a wide variety of such tools exist across sectors. Exemplary for the carbon composites industry, Rousseaux et al. analyzed over 600 tools, out of which

prominent examples are “Eco impact calculator” (online tool) or “Idemat database” (csv-based calculation sheet) [16, 18, 19]. Sector-specific tools require less extensive data, enabling quicker and simpler assessments. However, because their background databases are limited to specific sectors, they are unsuitable for cross-sectoral assessments but well-suited for evaluating standard use cases, particularly for SMEs [14–16].

In summary, a gap exists between the detailed coverage offered by commercial expert tools for cross-sector supply chains and the freely accessible, user-friendly online tools, which remain limited to specific sectors and lack multi-sector functionality.

2.3 Analysis of existing tools: GWP assessment in transportation

Assessing the GWP of transportation systems is essential for understanding their contribution to climate change and for identifying options to reduce greenhouse gas emissions across supply chains. A variety of analytical tools have been developed to support this task. These tools differ in scope, approach, and intended users, but they generally share a common purpose: to quantify the climate impacts associated with transport activities. The most used tools fall into two categories: general LCA software that can be adapted to model transport systems, and specialized transportation-focused tools designed specifically to estimate emissions from vehicle operation, fuel use, and transportation processes with varying levels of detail. Transportation-focused GWP assessment tools generally combine routing algorithms with mode-specific emission factors, load-dependent energy-consumption models, and detailed assumptions on vehicle technology and fuel pathways to estimate the environmental impacts of freight movements.

Schramm and Lehner *et al.* (2024) conducted a benchmarking study of online GHG emission calculators for freight transport, evaluating them in terms of transparency, completeness, and accuracy using a common benchmark shipment. Their analysis revealed that estimated GHG emissions for the same shipment can differ by more than a factor of two across tools, highlighting substantial methodological inconsistencies [20].

Among the general LCA tools that can incorporate transportation, the GREET Model offers a freely available platform for estimating greenhouse gas emissions across various transport technologies. It is provided both as an Excel-based implementation and as a .NET application with a graphical user interface [21, 22]. The tool includes a comprehensive life-cycle assessment approach that supports the evaluation of energy use and emissions across fuel-cycle pathways, as well as vehicle material production, assembly, operation, and end-of-life processes. While it is widely used in research and practice, the reliance on predominantly U.S.-based data sets, the level of methodological detail, the limited representation of transportation and infrastructure, and its standalone architecture may influence its suitability for non-U.S. settings or broader system-level analyses.

In contrast to general LCA tools, EcoTransIT World is a specialized tool designed specifically for assessing the environmental impacts of transportation activities. It provides detailed estimates of greenhouse gas emissions, energy consumption, and other pollutants across multiple transport modes, including road, rail, inland waterways, sea, and air freight. EcoTransIT World is available as an online web tool and provides a REST API, allowing integration into other software or custom applications for automated assessments. The tool is transparent about its data sources and calculation approach, as documented in its official methodology report, which details assumptions, emission factors, and modeling procedures used to estimate transport emissions [23].

2.4 Research gap

There is a growing variety of LCA models, databases, and approaches that support tailored assessments and wider stakeholder acceptance. However, many are sector- or region-specific and rely on fixed technology mixes and static emission factors, which limit their transferability and comparability across production processes, suppliers, and countries. Another limitation is that assessing production and transportation

separately carries a risk of double-counting or omissions. Additionally, SMEs face barriers to data availability and usability, so free tools often need to be combined. Finally, the changing energy systems over time and location, along with inconsistent handling of uncertainties, make it difficult to fairly compare procurement options and green energy scenarios. These gaps point to the need for a free, modular, and transparent approach that integrates production and transportation, supports data collection by SMEs, and accounts for uncertainties.

3. Methodological approach and calculation mode

3.1 General approach

The proposed approach aims to enable GWP calculation across multiple sectors by assessing emissions from both supplier-side production and transportation processes throughout the supply chain, supported by two dedicated steps. Firstly, the approach defines the supply chain by identifying all suppliers, including their locations and production types. Secondly, it allows the modelling of all production processes at each supplier tier, from Tier 1 to Tier n, together with the transportation steps connecting them along the supply chain. The data for this configuration step is derived from an LCI/LCA database, which stores either LCI data (relevant input and output flows, see Chapter 2.1) or LCA data (environmental impacts, see Chapter 2.1) for representative production and transportation processes, as well as environmental impacts of electrical and thermal energy supply data.

The database is based on openly accessible data sources, including literature, free LCA tools, and public LCI/LCA databases [24–26]. It includes two data types: process-specific LCI/LCA data sets, which provide GWP values based on the characteristics of each production process, and country-specific energy mix data, which reflect the GHG emissions of the energy source used, depending on location. In addition, it includes emission factors to calculate the GWP of each transportation process, considering the distance and transportation mode. Given the existence of adequate routing models, as previously referenced in Chapter 2, the routing functionality is not incorporated within the scope of this paper.

As the database is directly related to the supply chain configuration, it automatically links each process to the corresponding GWP value and subsequently calculates an overall GWP value for the entire supply chain by summing all individual GWP values for production and transportation into a single overall GWP score. The outlined approach is illustrated in Figure 3. This is carried out separately for each supply chain that is to be compared. Consequently, it gives the ability to identify emission hotspots and compare alternative configurations, such as different suppliers, production locations, or transport routes, providing a robust basis for informed and sustainable decision-making across the entire supply chain.

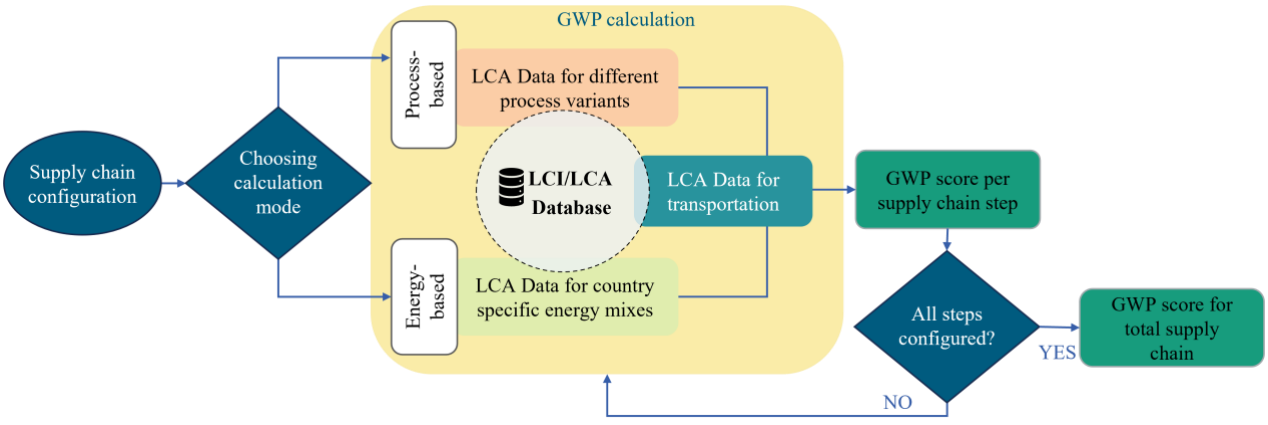


Figure 3: Visualization of the general approach, intended to calculate the GWP score.

For the GWP assessment of production steps, the explained approach offers two alternative calculation modes: process-based and energy-based. The process-based mode additionally provides a feature referred to as the black-box mode. The subsequent sections provide a detailed exposition of each of these modes.

The **process-based** mode enables comprehensive calculation using process-specific LCI/LCA data retrieved from the integrated database. These datasets typically include all relevant input and output flows associated with a process, such as materials, auxiliary substances, emissions, and energy use. As a result, this mode supports a detailed and complete GWP assessment of the supply chain. In addition, the process-based mode incorporates a fallback mechanism for supply chain stages where detailed process information is not available. In such cases, instead of relying on fully specified process data, the system substitutes these stages with representative or averaged LCA values obtained from the database, such as regional average emission factors when country-specific data is unavailable. This ensures that all process steps are assigned a GWP value, preserving the continuity of the assessment and enabling the calculation of a complete supply chain GWP score without compromising usability.

The **energy-based** mode, in contrast, focuses exclusively on the energy consumed during a process, distinguishing between thermal and electrical energy use. It does not consider additional inputs or outputs such as materials, emissions, or auxiliary substances. Its primary purpose is to enable comparisons of the energy-related aspects of different processes.

An important advantage of this mode is its applicability in situations where comprehensive LCA data is not available. As such, it is independent of any specific industrial sector and can be applied universally across diverse supply chain configurations. However, it requires direct input of the energy consumption for each process step. These values are then linked to the corresponding country-specific GWP factors, which are determined based on the defined process location.

Since the energy-based mode considers only energy consumption, while the process-based mode incorporates the full range of climate-relevant process flows, the results produced by the two modes differ fundamentally in scope. For this reason, the two calculation modes cannot be combined or directly compared, as doing so would lead to inconsistent or misleading interpretations.

3.2 Calculation mode

The calculation procedure implemented in the approach is designed to be straightforward and transparent.

In the process-based approach (Equation 1), the specific global warming potential of a process in the supply chain (GWP_{Pro}) is calculated by applying the GWP factor associated with the selected process variant ($GWPF_{ProV}$) and relating it to the mass of the product (m). These values are retrieved from the underlying LCI/LCA database and already include all relevant climate-related flows of the process. Consequently, no additional computation is required beyond selecting the relevant process variant.

$$\boldsymbol{GWP_{Pro} = GWPF_{ProV} \times m} \quad (1)$$

On the other hand, in the energy-based approach (Equation 2), the GWP of a process (GWP_E) is calculated as the product of the process's energy consumption (EC_{Pro}), the product mass (m), and the applicable country-specific GWP factor ($GWPF_E$).

$$\boldsymbol{GWP_E = GWPF_E * EC_{Pro} \times m} \quad (2)$$

Depending on the energy carrier, the national electrical or thermal energy mix is applied. For processes requiring both electrical and thermal energy, the respective contributions are calculated separately and then aggregated to obtain a single energy-based GWP value for the process.

Transportation emissions are calculated in a similar manner (Equation 3). For each route between two supply chain locations, the distance is determined. In the case of multimodal transport, intermodal terminals such

as ports or airports are included as intermediate points, and the distances for each mode and segment are evaluated separately. Subsequently, the transport-related GWP (GWP_{TR}) can then be assessed by multiplying these distances by the associated GWP factors ($GWPF_{TR}$).

In this way, each process step yields a specific GWP value that includes both production and transport contributions.

$$\boldsymbol{GWP_{Tr} = GWPF_{Tr} \times distance \times m} \quad (3)$$

Finally, the overall GWP of the supply chain (GWP_{SC}) is obtained by summing the GWP values of all individual process steps and transportation links (Equation 4). This aggregated score provides a comprehensive representation of the total climate impact of the configured supply chain and forms the basis for identifying hotspots. For each supply chain configuration, an aggregated GWP score is calculated and reported, enabling the comparison of alternative supply chain configurations with respect to their relative climate impacts.

$$\boldsymbol{GWP_{SC} = \sum GWP_{Pro} \ or \ \sum GWP_E + \sum GWP_{Tr}} \quad (4)$$

By allowing changes to specific elements of the supply chain, such as processes, supplier locations, and transport routes, the approach facilitates comparison of alternative supply chain scenarios and their effect on total GWP. This functionality provides decision-makers with data-driven insights to support more climate-friendly procurement and supply chain management strategies. For example, it can be used to evaluate how switching the energy source from a fossil-based supply to a renewable energy mix affects the overall GWP of the supply chain, or how changes in transport modes (e.g., from road to rail or from air freight to maritime shipping) influence the climate impact, thereby enabling comparison of alternative supply chain configurations.

4. Results and discussion

The following analysis presents a concrete example of greenhouse gas emissions within a composite sheet supply chain, illustrating how different production stages contribute to the overall carbon footprint (see Figure 4). The assessment uses process-specific GWP factors obtained from LCI and LCA databases, ensuring transparency and reliability in calculation (see Chapter 3.1). Tier 2 suppliers produce the raw materials: carbon fiber in China using conventional grid-mix energy, resulting in **2,046 kg CO_2-eq** or **60 kg** of carbon fiber (according to Equation 1), and Nylon PA6 in rope under conventional grid-mix energy, resulting in **450 kg CO_2-eq** or **40 kg** of resin. The transport of carbon fiber from China to Germany via air adds an additional **368 kg CO_2-eq**, while the transport of resin from Poland to Germany via road adds **4 kg CO_2-eq**. Tier 1 in Germany manufactures the composite sheet using grid-mix energy and conventional processes, contributing **67 kg CO_2-eq** or **100 kg** of product. The total supply chain generates up to **2,935 kg CO_2-eq** or **100 kg** of composite sheet. Normalizing these values or comparison shows that Tier 2 carbon fiber dominates the GWP contribution, accounting for roughly **70 %** of the total emissions, followed by Nylon PA6 at **15 %**, and Tier 1 composite manufacturing at only **2 %**. In the alternative scenario, the production of carbon fiber is relocated from China to Hungary, where a less carbon-intensive energy mix and more efficient production conditions reduce the GWO for carbon fiber production to **1,494 kg CO_2-eq** or **60 kg** of material, while the other stages remain unchanged. The transport of carbon fiber now occurs over a shorter distance by road, contributing only **5 kg CO_2-eq**, while the resin transport from Poland to Germany remains **4 kg CO_2-eq**. The Tier 1 composite sheet man act ring remains unchanged at **67 kg CO_2-eq**. The total supply chain GWP is therefore **2,021 kg CO_2-eq** or the **100 kg** composite sheet, representing a reduction of approximately **31%** compared to the original supply chain. This demonstrates that targeted improvements in both upstream material production and transportation logistics can significantly reduce the supply chain's overall carbon footprint, highlighting the potential for strategic interventions at the Tier 2 level.

As shown in the example above, the approach provides clear insights into the current supply chain and allows a transparent, straightforward assignment of alternative scenarios (different sectors, as well as including transportation and production processes). It helps quickly identify major contributors to greenhouse gas emissions and assess potential interventions.

Nevertheless, the approach has inherent limitations. The reliability of results depends strongly on the quality and completeness of LCI/LCA data, which can be particularly uncertain in specialized or niche domains. Moreover, relying on representative process values or average transport factors allows for quick, low-detail calculations but comes at the cost of reduced accuracy, potentially masking significant differences in supplier practices—such as variations in transport modes (ship versus air) or production technologies—making the method less appropriate for detailed, high-precision LCA analyses.

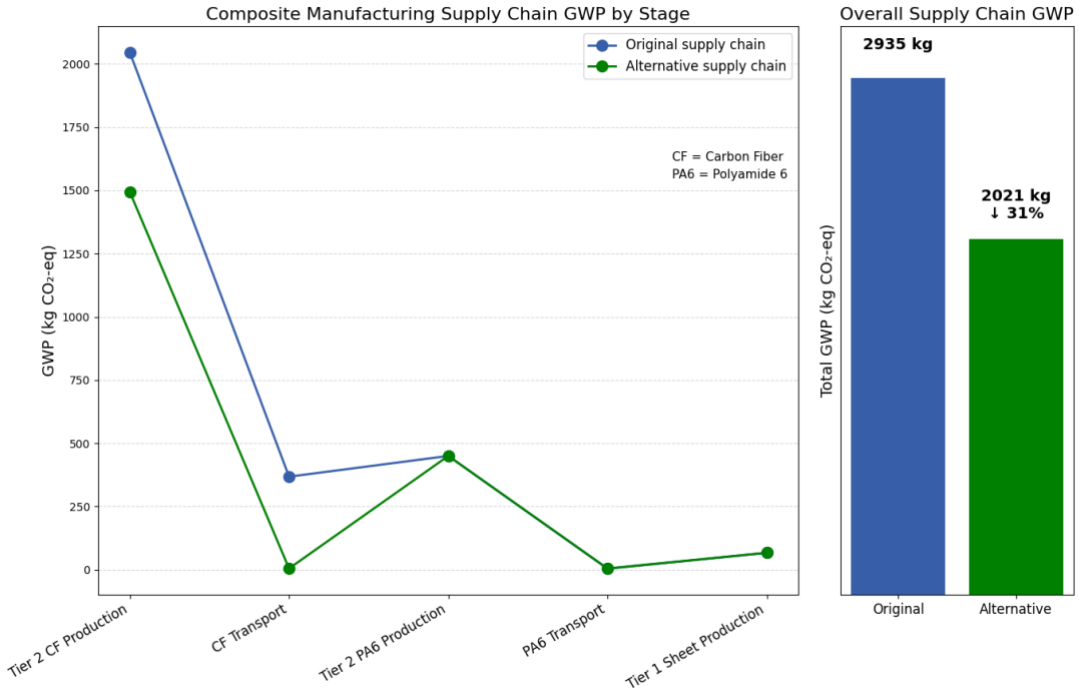


Figure 4: Application of the GWP assessment approach to a composite production supply chain, including the production and transport processes of carbon fiber (CF) and polyamide 6 (PA6) resin.

5. Conclusion and outlook

This paper presents an approach for assessing the GWP of production supply chains that is designed to be simple and easy to use, supporting evidence-based procurement and production decisions. By providing both a process-based and a simpler, energy-based screening mode, and by incorporating country-specific energy mixes and transport calculations, the approach is purposed to facilitate the identification of emission hotspots and the evaluation of alternative supplier, transportation, and energy scenarios. Although data quality and methodological consistency remain challenges, the approach is designed to be transparent and extensible, facilitating incremental improvements (through the collection of supplier data and database enhancements) and immediate use by SMEs as well as larger companies.

In summary, the proposed approach helps translate supply-chain emissions into actionable insights for sustainable procurement. Thereby, it identifies where to focus mitigation efforts, supports trade-off analyses, and strengthens the evidence base for optimized supplier selection. Future work will focus on validating the approach in real supply chains, quantifying uncertainty, and embedding the tool into procurement processes so environmental considerations become a routine element of supply-chain decision making.

Acknowledgment

This research was supported by the European-funded projects “ResC4EU – Resilient Supply Chains for Europe” (Project number: 101137643).

References

[1] United Nations Framework Convention on Climate Change (UNFCCC), 2015. Paris Agreement.

[2] Deng, Z., Zhu, B., Davis, S.J., Ciais, P. et al., 2025. Global carbon emissions and decarbonization in 2024 6, p. 231.

[3] Myhre, G., Shindell, D., Bréon, F.-M., Collins, W. et al., 2014. Anthropogenic and Natural Radiative Forcing, in Climate Change 2013 – The Physical Science Basis, Cambridge University Press, p. 659.

[4] Halat, K., Childe, S., Golding, N., Summerscales, J., 2026. Scope 3 Carbon Emissions Assessment for Manufacturing SMEs.

[5] Phillpotts, A., Owen, A., Norman, J., Trendl, A. et al., 2026. Completing the SME carbon profile: scalable prediction of scope 3 emissions.

[6] WRI & WBCSD, 2011. Greenhouse Gas Protocol: Corporate Value Chain Corporate Value Chain (Scope 3) Accounting and Reporting Standard.

[7] WRI & WBCSD, 2003. Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard Revised Edition.

[8] DIN Deutsches Institut für Normung e. V. DIN EN ISO 14044:2021-02, Umweltmanagement_- Ökobilanz_- Anforderungen und Anleitungen (ISO_14044:2006_+ Amd_1:2017_+ Amd_2:2020); Deutsche Fassung EN_ISO_14044:2006_+ A1:2018_+ A2:2020, Berlin. Beuth Verlag GmbH 13.020.10; 13.020.60.

[9] DIN Deutsches Institut für Normung e. V. DIN EN ISO 14040:2021-02, Umweltmanagement_- Ökobilanz_- Grundsätze und Rahmenbedingungen (ISO_14040:2006_+ Amd_1:2020); Deutsche Fassung EN_ISO_14040:2006_+ A1:2020, Berlin. Beuth Verlag GmbH 13.020.10; 13.020.60.

[10] Durão, V., Silvestre, J.D., Mateus, R., Brito, J. de, 2020. Assessment and communication of the environmental performance of construction products in Europe: Comparison between PEF and EN 15804 compliant EPD schemes 156, p. 104703.

[11] DIN EN ISO 14067:2019-02, Treibhausgase_- Carbon Footprint von Produkten_- Anforderungen an und Leitlinien für Quantifizierung (ISO_14067:2018); Deutsche und Englische Fassung EN_ISO_14067:2018, Berlin. Beuth Verlag GmbH 13.020.40.

[12] DIN EN ISO 14064-1:2019-06, Treibhausgase_- Teil_1: Spezifikation mit Anleitung zur quantitativen Bestimmung und Berichterstattung von Treibhausgasemissionen und Entzug von Treibhausgasen auf Organisationsebene (ISO_14064-1:2018); Deutsche und Englische Fassung EN_ISO_14064-1:2018, Berlin. DIN Media GmbH.

[13] Tubón-Núñez, E., Vigil Berrocal, M.Á., Rodríguez-Montequín, V., Ortega-Fernández, F., 2025. Tools to measure the environmental impact of a product 135, p. 1227.

[14] Adhikari, N., Li, H., Gopalakrishnan, B., 2025. A Bibliometric and Systematic Review of Carbon Footprint Tracking in Cross-Sector Industries: Emerging Tools and Technologies 17, p. 4205.

[15] Jusselme, J., Piaser, L., Bauer, T., 2025. Evaluating the carbon impacts of VSEs and SMEs: a simplified multi-sector tool through France and Italy 135, p. 51.

[16] Rousseaux, P., Gremy-Gros, C., Bonnin, M., Henriel-Ricordel, C. et al., 2017. “Eco-tool-seeker”: A new and unique business guide for choosing ecodesign tools 151, p. 546.

[17] UNFCCC Secretariat. The Greenhouse Gas (GHG) Emissions Calculator 2022. https://unfccc.int/documents/638361. Accessed 5 December 2025.

[18] European Composites Industry Association. Eco Impact Calculator for composites. https://ecocalculator.eucia.eu/Account/Login?ReturnUrl=%2F. Accessed 5 December 2025.

[19] Joost Vogtlander. Idemat2025RevA. https://www.ecocostsvalue.com/data-tools-books/. Accessed 5 December 2023.

[20] Schramm, H.-J., Lehner, M., 2024. Innovators and transformers: a benchmarking study of online carbon emission calculators for freight transport 54, p. 30.

[21] Wang, M., Elgowainy, A., Lee, U., Bafana, A., Banerjee, S., Benavides, P., Bobba, P., Burnham, A., Cai, H., Gracida, U., Hawkins, T., Iyer, R., Kelly, J., Kim, T., Kingsbury, K., Kwon, H., Li, Y., Liu, X., Lu, Z., Ou, L., Siddique, N., Sun, P., Vyawahare, P., Winjobi, O., Wu, M., Xu, H., Yoo, E., Zaimes, G., Zang, G., 2021. Greenhouse gases, Regulated Emissions, and Energy use in Technologies Model ® (2021 Excel). Argonne National Laboratory (ANL), Argonne, IL (United States).

[22] U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy. (2025). GREET Model: Greenhouse gases, Regulated Emissions, and Energy use in Technologies. Retrieved November 27, 2025, from https://www.energy.gov/eere/greet.

[23] EcoTransIT World Initiative. 2025. EcoTransIT World – Environmental Methodology and Data Update 2025. Retrieved from https://www.ecotransit.org (accessed 2025-11-27).

[24] Hannah Ritchie, Pablo Rosado, and Max Roser. Data Page: Lifecycle carbon intensity of electricity generation: Data adapted from Ember. https://archive.ourworldindata.org/20260304-094028/grapher/carbon-intensity-electricity.html. Accessed 20 August 2025.

[25] Climatiq Data Explorer. https://www.climatiq.io/data/explorer?sector=Energy&data_version=%5E24. Accessed 19 September 2025.

[26] Prenzel, T.M., Hohmann, A., Prescher, T., Angerer, K. et al., 2023. Bringing Light into the Dark-Overview of Environmental Impacts of Carbon Fiber Production and Potential Levers for Reduction. Polymers (Basel) 16.

Biography

Sönke Maatsch studied Economics at the Universities of Bremen (Germany), Lille 1 (France), and the New School (New York, USA). He completed his PhD in 2010. He is working as a research associate at the Institute of Shipping Economics and Logistics and is heading its Maritime Markets competence area. His research focuses on the economic impact, external costs and the resilience of maritime and multimodal transport chains.

Frida Schulze successfully completed her master's degree in biochemistry in 2024. Since 2024, she has been employed at the Fraunhofer IGCV, contributing to research and development activities in the field of sustainable production systems, with a particular focus on Life Cycle Assessment, specifically in the domain of composite production.

After graduating with a master`s degree in mechanical engineering, **Aljoscha Hieronymus** joined the MBtech Group in 2018 as a CAx consultant in development. He has been working at Fraunhofer IGCV since 2019 and has been leading the sustainable production systems team since 2024. His research focuses on planning and assessment of the ecological and economical sustainability of production systems for composites.

Clemens Gonnermann studied mechanical engineering at the Technical University of Munich (TUM), later focusing on process planning with an emphasis on assembly and industrial robotics. He completed his PhD at the Institute for Machine Tools and Industrial Management (iwb, TUM) on automated process monitoring and planning in assembly. In 2023, he founded the Department of Digital Production and AI at Fraunhofer IGCV (Institute for Casting, Composite Materials and Process Engineering), which he has headed ever since.

Oussama Etlijani studied Aerospace Engineering at FH Aachen and TU Braunschweig. Since 2022, he has been working at Fraunhofer IGCV, contributing to various research and industrial projects with a focus on IoT infrastructures and software development for production environments. His work focuses on bridging engineering and digitalization in the context of Industry 4.0.

Kunal Chaudhari completed his master's degree in Scientific Instrumentation at Ernst-Abbe-Hochschule Jena. Currently, he works as a research associate at Fraunhofer IGCV in the Smart Production and IoT team within the Department for Digital Production and AI, focusing on instrumentation, IoT solutions, and digitalisation for production environments.

Anja Munzke studied systems engineering and management with a specialization in logistics at the University of Applied Sciences THU in Ulm. 2019 she started working at the Fraunhofer Institute for Casting, Composite and Processing Technology (IGCV). Since 2024 Anja Munzke is head of department for sustainable factory planning and operation at the Fraunhofer IGCV.

Maximilian Holland heads the AI and Digital Engineering Lab at Fraunhofer IGCV, where he focuses on design automation through graph-based languages and the integration of generative AI into industrial manufacturing. His professional foundation bridges academic research at the University of Stuttgart and University of South Carolina with significant industrial experience at Premium AEROTEC / Airbus.

8th Conference on Production Systems and Logistics

Towards Principles for Designing Lean Intralogistics Systems: Systematic Transfer of Lean Thinking to Reduce Waste in Intralogistics

Lena Langenfeld[1], Thomas Bauernhansl[1,2]

[1]*Fraunhofer Institute für Manufacturing Engineering and Automation IPA, Stuttgart, Germany*
[2]*Institute for Industrial Manufacturing and Management (IFF), University of Stuttgart, Stuttgart, Germany*

Abstract

In lean production, the five well-known design principles of value specification, value stream orientation, flow, pull, and pursuit of perfection have proven successful [1]. However, these principles cannot be directly applied to intralogistics since logistics does not create value in the same manner as production. From the traditional perspective of lean production, large parts of logistics are considered wasteful. However, research and industrial practice have shown that well-designed intralogistics systems are essential for ensuring efficient and robust production processes [2].

The aim of this paper is, therefore, to develop principles for designing lean intralogistics systems that enable the systematic elimination of waste, thus supporting lean intralogistics goals. To this end, six types of waste specific to intralogistics are identified, with the key goals for intralogistics system design being discussed. Combining the waste and goals identified through literature on lean logistics and the principles of lean production, five system principles for lean intralogistics are derived: efficiency, holism, robustness, simplicity, and transparency and the overarching principle of effectiveness (i.e., user focus). Together with cultural principles like perfection and people focus, they form the basis for three design principles (process thinking, flow orientation, pull orientation). Expert interviews with industrial companies ensure the practical relevance and applicability of the results.

The principles form the novel basis for deriving specific design guidelines [3] that support the development of efficient, goal-oriented, and waste-free intralogistics systems - similar to the design guidelines of lean production, which have proven both effective and efficient in the value stream method [4].

Keywords

Design principles; Lean logistics; Lean intralogistics; Lean production; Intralogistics systems

1. From Lean Production to Lean Logistics: Requirements and Limitations

Since the introduction of the Toyota Production System (TPS) in the 20th century [5] lean production has become one of the most influential and widely adopted approaches to designing value-creation systems [6]. Its principles aim to eliminate waste by focusing on value-adding activities and continuous improvement [5]. Although lean production has been deeply researched and successfully applied for decades, its transfer to (intra-)logistics remains limited, as concepts cannot be transferred without adaptation. The reason lies in the fundamentally different nature of logistics, which is often insufficiently considered in traditional lean

DOI: https://doi.org/10.15488/21006
ISSN: 2701-6277

approaches. With logistics becoming increasingly important for creating customer value, its relevance has gained growing attention [7,8].

The fundamental distinction between production and logistics is how they generate value:
Production creates value through product transformation and should therefore align all activities with customer value, eliminating everything else as waste [1].
Logistics, in contrast, create value through utility in time and place and an increased service level [8,9]. Logistics comprises 'all operations needed to deliver goods or services, except making the goods or performing the services' [2]. Core logistics processes (transport, storage, handling, and order picking [10]) thus provide value by ensuring availability, correct location, timing, and presentation of material [2,8,11] – effectively bridging spatial and temporal gaps between the supply and the demand of material [8,10]. Intralogistics is, therefore, rather driven by *necessity* than by end-customer value (cf. Section 4).

To fully exploit the potential of the lean philosophy in intralogistics, and to systematically transfer and adapt concepts, methods, and tools from lean production, a comprehensive perspective on lean intralogistics is needed. This has not yet been achieved comprehensively:

- Some authors address types of waste in logistics, yet often without systematic differentiation or conceptual clarity [12,13]. In some cases, waste categories known from production are transferred directly without sufficiently considering logistics-specific characteristics [14]. A systematic and profound transfer from production to logistics is missing.
- Some authors examine lean for logistics in specific contexts like retail [12] or information logistics [15], but a holistic concept for lean in intralogistics, i.e., the logistics that covers internal material-handling and related operational processes [10], is missing:
  - Design and operation approaches mainly focus on the operational level and relate to specific design domains. Baudin [2], for instance, focuses primarily on transport and Kanban-based control, Jones et al. [16] discuss warehouse configuration, and Harris et al. [17] describe the practical use of part families, supermarkets, and milkruns for production supply.
  - Other authors particularly discuss the adaptation of individual lean methods to logistics, such as value stream mapping [18–20], which has proven highly effective in production design and optimisation [4]. However, fundamental theoretical considerations remain fragmented, preventing the systematic formulation of design principles for lean logistics.

As a result, a coherent conceptual foundation for a holistic lean intralogistics framework is still lacking. With such a framework being built, methods can systematically be transferred from production to logistics.

2. Research Question and Paper Structure

This paper aims to derive principles for designing lean intralogistics analogous to the established principles in lean production [1] to create this conceptual foundation. It addresses the **research question**: ***Which principles should guide the design of lean intralogistics?***
Such design-related principles require a clear understanding of overarching goals to unfold their instrumental value. Therefore, [sub-question (SQ) 1] *the goals that characterize lean intralogistics* must be identified.
The goal of every lean approach is to eliminate waste. Principles must be suited to support this. Therefore, it must be clarified, [SQ 2] *when intralogistics can be considered 'lean' and which types of waste must be eliminated*.
To finally derive principles, it must be defined [SQ 3] *what constitutes a principle and how it can be distinguished from related concepts*. This is especially necessary since concepts (e.g., principles and methods) are often mixed in literature and practice, as Choi et al. [21] also realized for the terms used in lean production.

After addressing SQ1, SQ2, and SQ3 in Sections 3, 4, and 5, Section 6 presents the literature-based derivation of principles for lean intralogistics to answer the research question. The derivation draws on established lean production principles as well as terminology describing 'good' logistics in literature. Expert interviews further support the results.

3. Goals of Lean Intralogistics

Just as in production and factory planning [22] the design of intralogistics systems is guided by a hierarchical set of goals. At the highest level, *formal goals* reflect a company's economic aims - such as maximising profit or profitability - to ensure long-term business success [23]. They are operationalized through *operational goals* [24], for example, the provision of specific goods.

From these emerge *performance goals* for the entire value-creation system, specifying *what* must be achieved. They include providing the right product, in the right quantity, at the right time, to the right place, and in the right quality. While production focuses on transforming inputs into outputs [25] intralogistics ensures that the required materials reach their point of use in production and that finished goods leave the factory. The performance goals of logistics are thus expressed through the *5Rs* (right material, right quantity, right time, right place, right quality) [18] and permeate the entire value chain.

On the next level, *design goals* specify *how* these performance goals should be pursued. The 'logical square of goals' [4] provides a framework distinguishing four non-substitutable and interdependent dimensions of design goals: Quality, speed, economy, and variability. Quality denotes the reliability and accuracy of achieving the 5Rs. Economy concerns the economic evaluation of logistics. Variability refers to the ability to adapt to changing requirements. Speed encompasses the rapidity of material flows and internal logistics processes. As the dimensions are mutually conflicting, they cannot be fully realised simultaneously. Consequently, the design of production and logistics systems requires a clear prioritisation among them. In intralogistics, this primarily manifests in balancing logistics costs (economy) with logistics performance (variability, quality, speed) [7,26]. The focus on these two dimensions reflects the perception of logistics in contrast to production: Processes are less technical than in production, and it is purely a question of cost whether performance goals can be achieved. Putting effort into making them lean and only implicitly more cost-efficient is pushed into the background easily.

Ideally, the strategic positioning is defined for both production and logistics at the factory level. In practice, production goals tend to be specified more explicitly than logistics goals. In such cases, logistics design should be aligned with the production design, since, e.g., a highly flexible intralogistics system adds limited value to a non-flexible production.

Principles and goals are strongly interrelated and need to be considered holistically. Fundamental principles should support the performance goal and apply equally to all design goals. Their validity does not depend on the prioritisation among design goals, nor do they influence it. Instead, they define the framework within which lean intralogistics must be designed and operated. In contrast, the application of design guidelines requires an explicit positioning within the conflicting design goals.
For lean intralogistics, this implies that principles must transcend individual design objectives and instead support the achievement of overall performance goals.

4. Waste in Lean Intralogistics

Everything that does not add value from the customer perspective prevents the achievement of goals and is defined as waste in lean production. Seven types of waste have traditionally been defined for production: overproduction, overprocessing, transport, inventory, unnecessary motion, waiting, and defects [1,5], supplemented by the waste of unused employee creativity as the eighth waste [27]. While the types of waste have been systematically analysed and transferred to information systems [15,28], their adaptation to lean

intralogistics is not yet systematically discussed - though mentioned [13,18]. Although the overarching idea of eliminating all activities that do not contribute value can be transferred to intralogistics, types of waste must be discussed and adapted due to fundamental differences in production and logistics (see Section 1). When distinguishing categories of waste, they should remain largely non-overlapping, meaning that one type of waste should not arise exclusively or automatically from another. Nevertheless, the defined waste categories do not claim to be complete or fully free from overlap. Having defined categories of waste should rather assist in identifying potentials systematically [18].

The performance goal of logistical processes is fulfilling the 5Rs (Section 3). Waste occurs either (1) when the 5Rs are not fulfilled or (2) when processes for achieving this goal are waste-full. According to [15], waste of category (1) is called primary waste, and waste of category (2) is secondary waste (Figure 1).

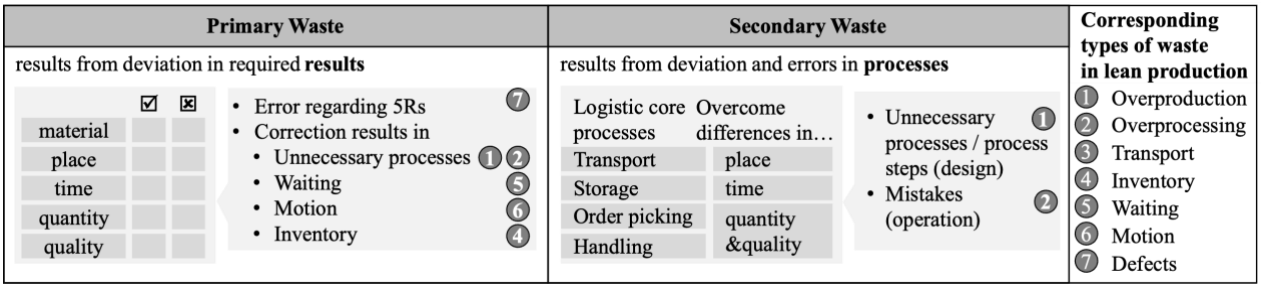


Figure 1: Waste in lean intralogistics

Primary waste occurs if an *error regarding the 5Rs* arises [29], when at least one of the 5Rs is not fulfilled, i.e., if the material itself, its quantity, the time, the place, or the quality of its presentation is not correct. In this case, originally *not necessary processes or process steps* have to be executed to correct the error: First, the process itself must be executed again to achieve a correct result, and second, the process must be reversed or corrected. If the wrong material is delivered to a workstation, it must be brought back to the warehouse or to the workstation where it is needed. *Waiting* occurs when the successor process cannot start because of missing material, and workers must wait. *Motion* occurs when processes must be corrected, and workers therefore move. *Inventory* is built up when material is stored because it is not in the correct place.

Secondary waste arises within logistics processes executed to meet performance goals. Such processes may generate waste without necessarily resulting in primary waste, since even inefficient processes may still produce correct outcomes. Transport, storage, order picking, and handling represent the core processes of intralogistics [10]. When required, they are not considered waste, but when unnecessary, they fall under *unnecessary processes*. Transport is required to compensate spatial, storage to compensate temporal, and handling and order picking to compensate qualitative discrepancies between supply and demand.

Determining the necessity of processes is complex and depends on the purpose of the process and the system objectives. Positioning within the design goals (Section 3) indicates the necessity: although core processes are non-value-adding, eliminating them would shift costs elsewhere, increase risk and planning effort, and reduce system flexibility [10]. Therefore, core processes are necessary because they are cost-efficient relative to transferring all tasks to the supply chain and help to stabilise the production by decoupling it from supply variability [2].

Reasonable inventory does not contradict just-in-time [21]. Necessary inventory is not waste [2,30,31], whereas excess inventory and unnecessary transport are, with unnecessary transport included in ‘unnecessary processes’. Inventory is addressed separately as it comprises storage inventory and process inventory, both of which can be identified through systematic analysis.

Logistic core processes are thus not questioned per se, yet their extent must be critically evaluated to determine the necessity of processes and process steps. Also, the way they are designed (e.g., sequence of steps, required material) can contain waste. Comparable findings for assembly processes are presented in [32], highlighting that process waste is difficult to detect and requires expert judgement.

Mistakes occurring during process execution are likewise considered waste. They tend to occur in complex processes with high variability and with many interfaces and branching points [33,34]. Like unnecessary processes, mistakes generate additional waste when they require correction. If such mistakes occur frequently, the underlying process design must be reconsidered.

The identified types of waste *(Errors regarding the 5Rs, unnecessary processes and process steps*, *mistakes*, *motion*, *waiting*, and *inventory*) correspond to waste in lean production as follows (see Figure 1):

- *Defects* (lean production) correspond to errors within the 5Rs, occurring whenever the actual outcome diverges from the target state.
- *Overproduction* and *overprocessing* (lean production) are reflected in unnecessary processes and process steps: activities are performed that are not required, either entirely or in part. *Transport* (lean production) is also subsumed under this category.
- *Waiting* and *Motion* apply to both production and intralogistics in an equivalent manner.
- *Inventory* is defined as waste in both lean production and lean intralogistics, although in intralogistics it is considered waste also insofar as it becomes visible without being directly attributable to intralogistics itself.

5. Principles for System Design

Design principles are the framework that guides system design and helps to achieve waste-free and goal-oriented systems. They do not provide concrete instructions for designing a system, but the framework within which one can operate. They indicate the direction one should follow [3]. Principles are enduring, universal, and abstract. As a normative instance, they help to identify the most suitable solution among possible alternatives [35]. They apply independently of specific technologies or industries and must be interpreted and applied in a context-specific manner. Their formulation must be clear and unambiguous. Orientation towards principles supports the achievement of overarching goals. Within a set of principles, all elements are equal and internally consistent, i.e., they do not contradict each other. When combined into a set, principles do not have to be totally mutually exclusive, but they should not be completely substitutable or dependent. Requirements for principles are summarized in Figure 2a.

Principles must be distinguished from design guidelines and design rules that can be deduced from principles [3]. Erlach et al. [3] arrange principles, guidelines, and rules in a triangle to reflect the hierarchy and number of principles (few) over guidelines (some) to rules (many) (Figure 2b). Furthermore, principles must not be confused with methods (i.e., standardized and stepwise procedure descriptions) or goals (i.e., desired end states or outcomes). For clarity about the terminology used and to disentangle terminology often mixed in the literature [18,36,37] Figure 2b shows the definitions used in this paper.

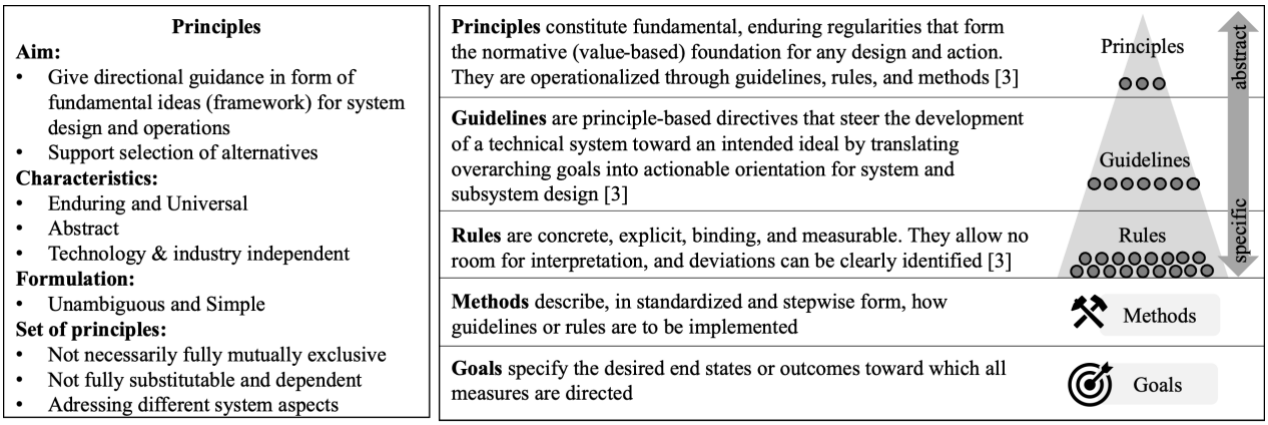


Figure 2a (left): Description of principles; Figure 2b (right): terminology and hierarchy (adapted from [2])

Principles often emerge empirically, for example, through systematic observation, experiments, or studies that reveal which procedures lead to successful outcomes. Cultural and historical influences also shape their formation. For instance, Taiichi Ōno drew the lean philosophy not only on observations of American supermarkets but also on Japanese values of respect, order, and team orientation, shaped by post-war resource scarcity [5]. Based on practical observations, he derived the core elements of the TPS. Empirical studies [6] and theoretical extensions reinforced and integrated these elements into the broader framework of lean production.

The principles *Specify Value*, *Identify the Value Stream*, *Flow*, *Pull*, and *Pursue Perfection* have traditionally guided lean production systems [1]. However, Erlach et al. correctly note that these do not consistently constitute fundamental design principles formally, as some of them prescribe concrete actions [3]. They therefore propose 'preliminarily and tentatively' *customer-oriented*, *waste-free*, and *defect-free* as principles of lean production [3]. Based on this, and acknowledging that traditional lean principles have proven effective over many years and have, through their continued use and dissemination, evolved into principles regardless of formal definition, this paper distinguishes between *system principles* and *design principles*.

System principles are superordinate, broadly understandable concepts, independent of lean terminology. They describe the required design characteristics of an intralogistics system to ensure reliable, economically efficient, and high-performance outcomes. For production, they compromise *customer-oriented*, *waste-free*, and *defect-free* [3]. These principles apply to every design or operation and are the basis for design principles.

Design principles are specific to the lean context and specify the purpose and perspective of the system as well as functional and structural logic. They must refer to the system as a whole, not to specific subsystems. *Specify Value, Identify the Value Stream, Flow,* and *Pull* are design principles.

Those principles are supported by **cultural principles**, that build the fundamental attitude of lean organizations and shape behaviour and decision-making. Instead of prescribing how systems should be designed, they create the conditions under which system and design principles can be effectively applied and sustained over time. *Pursuit of perfection* [1,30] can be classified as a cultural principle. It refers to the cultural foundation of continuous, incremental improvement aimed at the systematic elimination of waste. *People focus* [18,38] emphasizes the socio-technical nature of factories, acknowledging that their performance depends not only on technical design but also on human involvement.

6. Principles in Lean Intralogistics

6.1 Approach to Derive Principles for Lean Intralogistics

Principles of lean intralogistics are derived through a four-step approach:

1. First, aspects that have been referred to as 'principles' or similar ('basic principles' [37], 'design principles' [38], 'guidelines' [18]) in literature about lean production [1,5,38–40] and lean (information)logistics [2,15,16,29,36,37,41–43] are examined. Appendix A shows the results.
2. To broaden the perspective, further literature [14,26,43–48] is reviewed for often-used but not explicitly denoted principles that describe 'good' logistic systems. Appendix B shows the results.
3. Expert interviews supplement the literature findings with practical insights. The sample consisted of seven representatives from small and medium-sized companies in southern Germany, covering mechanical and plant engineering, automation technology, and tooling manufacturing. The results are consolidated in a collaborative workshop and are documented in Appendix C.
4. The identified aspects are evaluated to determine whether they (a) meet the requirements for principles (Figure 2a), (b) apply for the system as a whole and not for subsystems and can be distinguished from other concepts (Figure 2b), (c) complement each other effectively without being fully substitutable or dependent on each other. In case of high overlap or dependencies, results were combined in one cluster.

The approach yielded the following main findings:

(1) *The principles of lean production were often transferred without justification and expanded to include various additional aspects without engagement with the concept of 'principles' itself.*
Classical lean principles have largely been transferred to the domain of lean logistics [13, 18,36,37,49]. They have been extended by concepts referred to as 'principles' that incorporate specific technical solutions (e.g., flow racks [40]) or implementation approaches (e.g., milkruns [40]) as well as design approaches (e.g. line-back planning [36,37]) or planning approaches (e.g. logistics levelling, [36]) intended to support lean goals. Furthermore, approaches that may be classified as design guidelines, can be found (e.g., standardisation, synchronization, takt, integration [13,18,37,49]). The coexistence of principles and closely related concepts may be attributed to the absence of a clear and consistent definition of what constitutes a principle. A systematic discussion of the suitability and transferability of production-oriented principles to the logistics context remains largely absent. Similar to lean production, the principles presented are primarily design-oriented and can therefore be positioned within the broader layer of design principles.

(2) *Different levels and varying degrees of detail allow for hierarchical structuring and clustering.* Several of the identified principles can be subsumed under broader principles. For example, stability- understood as robustness at a specific point in time - was integrated into the overarching concept of robustness. Especially when evaluating the different terms that describe 'good' logistic systems, many terms were used that describe similar concepts.

6.2 System and Design Principles in Lean Intralogistics

Figure 3 shows the resulting principles. They are detailed and explained in the following section.

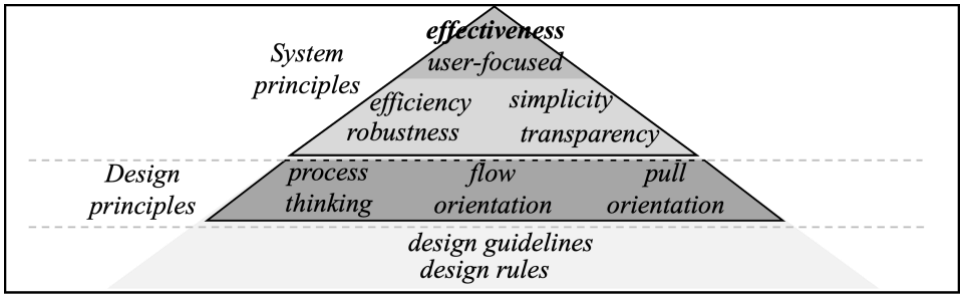


Figure 3: Principles for lean intralogistics

Five **system principles** for design and operations were derived, with *effectiveness* being superior to *efficiency, robustness, simplicity,* and *transparency.*

Effectiveness extends the customer orientation of lean production to intralogistics, where the 'customer' includes not only external recipients but also internal users, most importantly production, which needs a reliable material supply. The neutral term 'user' replaces 'customer' (c.f. [15]). Creating value for all users and fulfilling their requirements is the primary objective of lean systems [4,5]. Accordingly, all design and operational decisions must be aligned with user needs. Effectiveness should be addressed before any further aspect [2] and forms the foundation of lean intralogistics as overarching principles. All authors focusing on lean logistics agree on this, though its distinction from production are scarcely discussed.

Efficiency [2] implies achieving the intended output with minimal input, not only explicitly from a cost perspective. Waste impedes this: systems containing waste inevitably require additional effort. For example, executing unnecessary processes or process steps consumes time and resources, reducing overall efficiency. Mistakes reduce the efficiency of the system, as do unnecessary motion and waiting.

Robustness [13] is characterised by stability and reliability under variable conditions, resistance to disturbances, and the flexibility to adapt to change while maintaining core functionality. Robust processes are reliable, predictable, reproducible, and designed to minimise errors.

Simplicity [30], meaning a low level of complexity, helps to achieve stability. Standardisation can support this, but only if applied purposefully and in combination with continuous improvement to prevent stagnation. Simplicity does not only allow for robustness, but also for speed and low cost.

Transparency [e.g., 33,37] in systems makes the current system state and relevant information visible, enabling the timely identification of deviations and inefficiencies. Transparency makes all types of waste visible and allows for finding the root causes to be able to actively work on them.

Process thinking, *flow orientation*, and *pull orientation* serve as **design principles**. They represent three complementary perspectives on lean intralogistics design: Process thinking as *structural perspective* focuses structuring and standardizing end-to-end activities, flow as *system behaviour perspective* emphasizes the continuous and smooth material and information flow. Pull as *control perspective* addresses the demand-driven triggering and control of logistics operations based on actual consumption.

Process thinking refers to viewing material and information flows as interconnected end-to-end processes rather than isolated functions, aiming to optimise overall flow, reduce interfaces, and improve performance across the entire chain starting from the user's perspective. In combination with the principle of effectiveness, process thinking corresponds to 'value stream orientation' in lean production. Achieving genuine end-user orientation requires a holistic view of intralogistics from in- to outbound logistics [13,50]. Approaches from general process design [33,34] can be used design and optimize intralogistics.

The *flow principle* [50] aims to establish uninterrupted material movement across all process steps. Material in flow does not stop without a reason, accumulate in buffers, or incur waiting times. Continuous flow enhances process stability and predictability. The principle applies equally to material and information flows, as both must progress synchronously to maintain system efficiency.

The *pull principle* ensures that material is moved only when demanded by the customer. In this sense, it directly operationalises user focus. It positively affects inventory performance, reduces and stabilizes lead time, does it almost automatically, is suited for almost any production system, and is robust [51]. By preventing premature or unnecessary processes, a pull system supports the 5Rs and reduces inefficiencies.

7. Conclusion and Outlook

This paper develops a conceptual foundation for designing lean intralogistics by deriving both system principles and design principles that are analogous to, yet distinct from, those used in lean production. Building on the observation that intralogistics creates value differently from production and is rather driven by the necessity of processes than by end-customer value, the paper first shows the structure of goals in intralogistics and concludes that principles must support the performance goals but may not restrict the specification of design goals. Furthermore, *errors in 5Rs, unnecessary processes and process steps, motion, waiting, inventory*, and *mistakes* are proposed types of waste in intralogistics.

Linking these waste categories with the overarching design and performance goals of intralogistics enables a structured derivation of principles. The resulting system principles (*effectiveness, efficiency, holism, simplicity, robustness, transparency*) describe the essential characteristics of lean intralogistics systems and provide a normative orientation for system design and operation. The design principles (*process thinking, flow orientation, pull orientation*) translate this orientation into guiding ideas that shape functional and structural logic. Expert interviews confirmed the practical relevance of these principles and highlighted their applicability across different contexts. Cultural principles support design and system principles.

The results offer a conceptual framework that allows lean methods and tools to be transferred from production to logistics in a systematic manner. They serve as a basis for the development of design guidelines that support practitioners in analysing, designing, and optimizing intralogistics as they do in production [3,4]. Further research and application in practice may focus their development and prove their value in practice.

Acknowledgements

The project is funded by the Federal Ministry for Economic Affairs and Energy based on a resolution passed by the German Bundestag.

Appendix A. Principles in literature

Table 1: Lean principles in literature

Principle in Literature	Author(s)	Comment & Classification
Value orientation (from the customer's perspective)	[1,13,18]	***Overarching system principle:*** Philosophical, strategical principle for describing the purpose of any action: Why is any action being taken? → General rule for when a process adds value - Interpretation depends on the use case. **But:** *"Value" hast to be redefined in logistics, "customer" is user.*
user orientation	[15]	
Value stream orientation	[1,13,18,38,39]	***System principle:*** Thinking principle: Thinking in terms of processes, not functional areas *"Value stream" and "waste" have to be redefined in logistics* → Becomes **process thinking**
Purposeful	[15]	
Flow	[1,15,36,37,39,49]	***Design principle: Flow principle:*** Systemic principle: Describing a system characteristic and methodological orientation
Pull	[1,13,15,18,36,37,39,49]	***Design principle: Pull principle:*** Organisational principle: Defines the functional logic of the control system. Abstract, timeless, value-oriented, various methods are related to the principle (e.g., Kanban, CONWIP)
Standardization	[13,18,37–39,49]	*Design guideline:* **Approach and basis for pull and flow;** Standardization must be applied goal-oriented and purposeful where it is useful and dependent on the system goals → not valid as principle
Synchronisation	[36,37,49]	*Design guideline:* **Implementation element** / concept for implementing flow principle
Takt	[37,39,49]	Implementation approach for achieving pull/flow (cycle stabilizes flow)
Pursue Perfection	[1,13,15,18,38]	***Cultural principle*** of learning, continuous improvement, error culture. results in the design of an ideal system state that can be achieved in several steps
Zero Defect principle	[38]	
Employee orientation	[18,38]	***Cultural principle*** Valid principle, but not in focus or process design / optimization
Robustness	[13,18]	***System principle***
Long-Term perspective		Included in robustness: Robustness means applicability in different settings and therefore in long-term perspective
Holism		Planning approach
Stability	[37,49]	Included in robustness: Robustness means stability in different settings
Integration		*Design guideline:* "Integrate process steps if possible"
Transparency	[36]	***System principle***
Flexibility		*Included in / subordinate to robustness: to achieve robustness, a system has to be flexible*
Logistics Levelling		Method
Logistics-oriented product design		concerns department of product design, not logistics
Line-Back-Principle		Planning principle included in process thinking and user-focus
Integrated storage systems	[40]	Logistics technology & solution approach for subsystem
Conveyor belt-based supermarket		Logistics technology & solution approach for subsystem
Stop optimization		Solution approach for subsystem
Crossing-free traffic		Solution approach for subsystem
Takted milkrun		Solution approach for subsystem
Forklift-free logistics		Solution approach for subsystem
Intelligent container management		Logistics technology
Flow racks		Logistics technology
Line Runner		Method (to achieve levelled and stable systems)

Bold: traditional lean principles

Appendix B: Characteristics of 'good' logistics processes in literature

Table 2: Characteristics of 'good' and lean logistics in literature

	General Literature about Logistics						Literature about Lean Logistics							
Term **Author**	**[26]**	**[14]**	**[46]**	**[48]**	**[52]**	**[8]**	**[40]**	**[15]**	**[36]**	**[53]**	**[41]**	**[18]**	**[37]**	**[42]**
effective (user-/ customer-focused)	x	x	x			x	x	x	x	x	x	x	x	x
targeted, goal-oriented	x	x	x								x	x		
impactful				x										x
customer-oriented			x	x		x	x	x	x	x	x	x	x	x
use-and service-oriented					x	x		x						
efficient	x	x	x	x		x	x	x	x		x	x	x	x
cost-effective	x		x	x			x			x	x			x
fast (time-efficient)		x	x	x	x	x	x	x	x	x	x	x	x	
responsive	x		x		x	x					x	x		
energy-saving					x									
error-free					x					x	x	x		
holistic		x	x	x	x	x	x		x	x	x	x	x	x
comprehensive		x	x	x		x	x		x	x	x		x	x
integrative		x	x			x	x		x				x	
interconnected		x	x			x		x			x			
cross-departmental		x		x	x	x	x				x			
robust		x	x				x			x	x	x	x	x
stable		x	x	x			x	x	x	x	x	x	x	
reproducible		x	x											
interruption free	x			x		x								x
standardized	x		x		x		x		x	x	x		x	x
predictable										x		x		x
simple	x	x	x	x	x	x	x		x	x	x	x	x	x
intuitive		x		x								x		
structured				x	x		x	x	x	x	x	x	x	x
transparent			x		x		x	x		x	x	x	x	

x *term used to describe logistics*

Appendix C: Expert interview

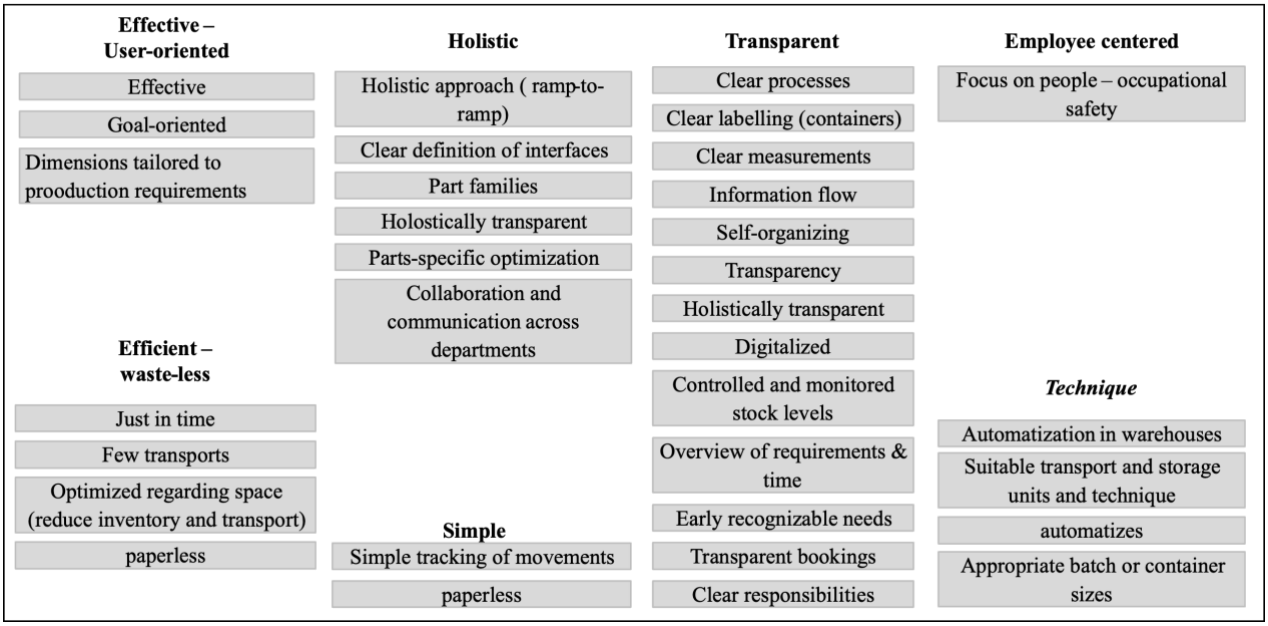


Figure 4: Results of expert interviews about 'good' lean intralogistics

References

[1] Womack, J.P., Jones, D.T., 2003. Lean thinking: Banish waste and create wealth in your corporation, Rev. and updated, 1. paperback ed. Simon & Schuster, London.

[2] Baudin, M., 2005. Lean logistics: The nuts and bolts of delivering materials and goods. Productivity Press, New York, NY.

[3] Erlach, K., Böhm, M., Hartleif, S., Leipoldt, C., Ungern-Sternberg, R., 2020. Gestaltungsrichtlinien in den Technikwissenschaften. Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (1-2), 77–81.

[4] Erlach, K., 2013. Value stream design: The way towards a lean factory, 1st ed. Springer, New York.

[5] Ōno, T., 2008. Toyota production system: Beyond large-scale production, [Reprinted] ed. Productivity Press, New York, NY.

[6] Womack, J.P., Jones, D.T., 1990. The Machine That Changed the World: The Story of Lean Production. Free Press, New York.

[7] Schulte, C., 2013. Logistik: Wege zur Optimierung der Supply Chain, 6., überarbeitete und erw. Aufl (Online-Ausg.) ed. Verlag Franz Vahlen, München.

[8] Pfohl, H.-C., 2018. Logistiksysteme: Betriebswirtschaftliche Grundlagen, 9. Aufl. 2018 ed. Springer, Berlin, Heidelberg.

[9] Göpfert, I., Braun, D., Schulz, M. (Eds.), 2017. Automobillogistik: Stand und Zukunftstrends, 3rd ed. Springer Gabler, Wiesbaden.

[10] Heiserich, O.-E., 2011. Logistik: Eine praxisorientierte Einführung, 4., vollständig überarbeitete und erweiterte Auflage ed. Gabler Verlag, Wiesbaden.

[11] Huppertz, A., Stühlen, M., 2013. Lean Logistics – Made by Huppertz, in: Göpfert, I. (Ed.), Automobillogistik. Stand und Zukunftstrends, 2., aktualisierte und erw. Aufl. ed. Springer Gabler, Wiesbaden, pp. 305–329.

[12] Macedo, H., Tomaz, L., Guimarães, L., Cerqueira-Pinto, L., Sá, J.C., Dinis-Carvalho, J., 2025. Development of a Classification Model for Value-Added and Non-Value-Added Operations in Retail Logistics: Insights from a Supermarket Case Study. Sustainability 17 (7), 3177.

[13] Durchholz, J., 2013. Leitlinien für eine schlanke Logistik, in: Günthner, W.A., Durchholz, J., Klenk, E., Boppert, J. (Eds.), Schlanke Logistikprozesse. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 9–15.

[14] Muchna, C., Brandenburg, H., Fottner, J., Gutermuth, J., 2018. Grundlagen der Logistik: Begriffe, Strukturen und Prozesse. Springer Gabler, Wiesbaden.

[15] Hartleif, S.M., 2022. Gestaltung einer schlanken Informationslogistik im komplexen Produktionsumfeld. Stuttgart: Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA.

[16] Jones, D.T., Hines, P., Rich, N., 1997. Lean logistics. International Journal of Physical Distribution & Logistics Management 27 (3/4), 153–173.

[17] Harris, R., Harris, C., Wilson, E., 2011. Making materials flow: A lean material-handling guide for operations, production-control, and engineering professionals, Version 1.1, July 2011 ed. Lean Enterprise Institute, Cambridge, Mass.

[18] Durchholz, J., 2014. Vorgehen zur Planung eines schlanken Logistikprozesses: Wertstromdesign für die Logistik. Zugl.: München, Techn. Univ., Fak. für Maschinenwesen, Diss., 2014. Techn. Univ. Lehrstuhl für Fördertechnik Materialfluss Logistik fml, München.

[19] Knoll, D.N.G., 2022. Value Stream Mapping for Internal Logistics using Process Mining. utzverlag, München.

[20] Balsliemke, F., 2015. Kostenorientierte Wertstromplanung: Prozessoptimierung in Produktion und Logistik. Springer Gabler, Wiesbaden.

[21] Choi, T.Y., Netland, T., Sanders, N., Sodhi, M.S., Wagner, S.M., 2023. Just-in-Time for Supply Chains in Turbulent Times. SSRN Journal.

[22] Erlach, K., Berchtold, M.-A., Kaucher, C., Kronschnabl, A., 2025. Systematische Festlegung von Fabrikzielen. Zeitschrift für wirtschaftlichen Fabrikbetrieb 120 (7-8), 489–495.

[23] Thommen, J.-P., Achleitner, A.-K., Gilbert, D.U., Hachmeister, D., Jarchow, S., Kaiser, G., 2023. Allgemeine Betriebswirtschaftslehre, 10., überarbeitete und aktualisierte Auflage ed. Springer Gabler, Wiesbaden.

[24] Wöhe, G., 2010. Einführung in die allgemeine Betriebswirtschaftslehre, 24th ed. Vahlen, München.

[25] Günther, H.-O., Tempelmeier, H., 2012. Produktion und Logistik, 9th ed. Springer, Berlin.

[26] Gudehus, T., 2011. Logistik: Grundlagen - Strategien - Anwendungen, 4., aktualisierte Aufl 2010 ed. Springer-Verlag, Berlin, Heidelberg.

[27] Liker, J.K., 2004. Toyota way: 14 management principles from the world's greatest manufacturer, First edition ed. McGraw-Hill Education; McGraw Hill, New York, N.Y.

[28] Meudt, T., Leipoldt, C., Metternich, J., 2016. Der neue Blick auf Verschwendungen im Kontext von Industrie 4.0. Zeitschrift für wirtschaftlichen Fabrikbetrieb 111 (11), 754–758.

[29] Günthner, W.A., Durchholz, J., Klenk, E., Boppert, J. (Eds.), 2013. Schlanke Logistikprozesse. Springer Berlin Heidelberg, Berlin, Heidelberg.

[30] Frazelle, E., Ward, R., Apple, M, J., Day, T., Petrina, G., Voss, A., Zollinger, H., 1993. JTEC panel report on material handling technologies in Japan.

[31] Frazelle, E., 2023. World-class material handling and warehousing. McGraw-Hill Education, New York.

[32] Baudin, M., 2006. Lean assembly: The nuts and bolts of making assembly operations flow, [Nachdr.] ed. Productivity Press, New York, NY.

[33] Best, E., Weth, M., 2010. Process Excellence: Praxisleitfaden für erfolgreiches Prozessmanagement, 4., überarb. u. erw. Auflage ed. Gabler, Wiesbaden.

[34] Schuh, G., 2006. Change Management: Prozesse strategiekonform gestalten. Springer, Berlin, Heidelberg, 123 pp.

[35] Kesselring, F., 1954. Technische Kompositionslehre: Anleitung zu technisch-wirtschaftlichem und verantwortungsbewußtem Schaffen. Springer, Berlin, Heidelberg.

[36] Hoeltz, N. Lean Logistics Maturity Model: ein Reifegradmodell zur Bewertung schlanker intralogistischer Unternehmensstrukturen.

[37] Klug, F., 2018. Logistikmanagement in der Automobilindustrie: Grundlagen der Logistik Im Automobilbau, 2nd ed. ed. Springer Berlin / Heidelberg, Berlin, Heidelberg, 1528 pp.

[38] VDI - Verein deutscher Ingenieure, 2012. Ganzheitliche Produktionssysteme: Grundlagen, Einführung und Bewertung.

[39] Hänggi, R., Fimpel, A., Siegenthaler, R., 2021. LEAN Production – einfach und umfassend: Ein praxisorientierter Leitfaden zu schlanken Prozessen mit Bildern erklärt. Springer Vieweg, Berlin, Heidelberg.

[40] Schneider, M., 2021. Lean Factory Design: Gestaltungsprinzipien für die perfekte Produktion und Logistik, 2nd ed. Hanser, München.

[41] Günthner, W.A., Boppert, J. (Eds.), 2013. Lean Logistics: Methodisches Vorgehen und Praktische Anwendung in der Automobilindustrie, 1st ed. ed. Springer, Berlin, Heidelberg.

[42] Goldsby, T.J., Martichenko, R., 2005. Lean Six Sigma logistics: Strategic development to operational success. J. Ross Publ, Boca Raton, Fl.

[43] Schmidt, T. (Ed.), 2019. Innerbetriebliche Logistik. Springer Vieweg, Berlin.

[44] Furmans, K., Kilger, C. (Eds.), 2018. Gestaltung der Struktur von Logistiksystemen, 1st ed. Springer Berlin Heidelberg, Berlin, Heidelberg.

[45] Furmans, K., 2019. Betrieb Von Logistiksystemen. Springer Berlin / Heidelberg, Berlin, Heidelberg.

[46] Lasch, R., 2018. Strategisches und operatives Logistikmanagement: Prozesse, 2., überarbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden.

[47] Schulz, S., Technische Universität Ilmenau, 2011. Gestaltung von Logistiksystemen in der Fertigungsindustrie: ein typologisch orientierter Ansatz. Ilmenau.

[48] Arndt, H., 2015. Logistikmanagement. Springer Gabler, Wiesbaden.

[49] Klug, F., 2008. Gestaltungsprinzipien einer Schlanken Logistik. Zeitschrift für die gesamte Wertschöpfungskette Automobilwirtschaft, 56–61.

[50] Wegner, U., 2017. Einführung in das Logistik-Management: Prozesse – Strukturen – Anwendungen, 3rd ed. Springer Gabler, Wiesbaden.

[51] Roser, C., Shook, J., 2021. All about pull production: Designing, implementing, and maintaining Kanban, CONWIP, and other pull systems in Lean Production. AllAboutLean.com Publishing, Offenbach.

[52] Martin, H., 2016. Transport- und Lagerlogistik: Systematik, Planung, Einsatz und Wirtschaftlichkeit, 10. Aufl. 2017 ed. Springer Vieweg, Wiesbaden.

[53] Günthner, W.A., Durchholz, J., Klenk, E.M., Boppert, J., Knössl, T., Klevers, M., 2013. Schlanke Logistikprozesse: Handbuch für den Planer. Springer Berlin Heidelberg, Berlin, Heidelberg.

Biography

Lena Langenfeld (*1992) is a research assistant in the group Operational Excellence at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart, Germany. Since 2021 she has been working as a research consultant at the Fraunhofer IPA with a focus on production optimisation and factory and logistics planning, in both industry and research projects.

Prof. Dr.-Ing. Thomas Bauernhansl (*1969) has been head of the Fraunhofer IPA in Stuttgart and the IFF at the University of Stuttgart since September 2011. Bauernhansl studied mechanical engineering at the RWTH Aachen and received his doctorate in 2002. He then worked for Freudenberg, most recently as Director of the Technology Centre at Freudenberg Sealing Technologies. His scientific focus areas include personalised production, biological transformation, and Industry 4.0.

8th Conference on Production Systems and Logistics

Development Of A Methodology To Determine Change Variance In Product Architecture Elements In The Context Of The Upgrade Circular Economy

Günther Schuh[1,2], Alexander Keuper[1], Frederike Hellwig[1] , Kira Rimbach[1]

[1] *Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Aachen, Germany*
[2]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

Increasing resource scarcity, environmental effects and rising regulatory pressure require a transition from linear economy towards upgrade circular economy. Within this context, re-assembly and upgrade strategies aim to extend product lifecycles by enabling value-adding improvements over multiple use phases. However, the successful implementation of such circular approaches is often hindered by a lack of transparency regarding where and how changes may occur within a product architecture. Dynamic influences from markets, customer needs, and technological developments are difficult to anticipate, leading to inflexible designs and missed upgrade opportunities. This paper introduces a concept for a methodology to determine the change variance of product architecture elements in the context of the circular economy. The approach combines foresight-based analyses with architectural modelling to identify and quantify the potential of future changes. By systematically linking market and technology trends to functional and structural product representations, it becomes possible to locate those elements most susceptible to change and assess their impact on re-assembly and upgrade potential. The resulting evaluation of architectural flexibility and criticality supports designers in developing future-resilient architectures that can accommodate evolving requirements. The proposed methodology enables manufacturing companies to proactively integrate expected change into early product development stages. It thereby enhances upgradeability, extends product lifetime, and fosters the practical realization of circular economy principles through adaptive and foresight-driven product architectures.

Keywords

Upgrade Circular Economy; Upgrades; Foresight; Change Variance; Product Architecture

1. Introduction

In the manufacturing industry, sustainability is evolving from an optional guiding principle into a central necessity [1]. Companies face the challenge of meeting the rising demands on technical products while remaining competitive [2]. The upgrade circular economy (UCE) decouples resource use from economic growth and offers a resource-efficient alternative for improving products, with value creation realized through product upgrades [3]. One UCE concept that implements the sustainable, industrial reconditioning of products is re-assembly [4]. In order to provide an appropriate amount of flexibility and adaptability for re-assembly [5], future-proof product architectures (PAs) with these characteristics must be developed [6]. A key success factor here is the ability to respond to dynamic customer and market requirements and integrate late changes into the product architecture [6].

DOI: https://doi.org/10.15488/21007
ISSN: 2701-6277

Product upgrades help accommodate changes by making products and their PAs more sustainable and more adaptable. This requires the early identification and consideration of future requirements, particularly in modular architecture with long-lived elements [7]. Foresight enables a holistic analysis of trends, technologies, and influencing factors, thereby creating the foundation for integrating changing requirements into product design [8]. By applying foresight continuously, uncertainties can be reduced, and robust, sustainable PAs can be created that facilitate re-assembly processes. As a first step towards sustainable PA, this paper develops a methodology to systematically identify the change variance (CV) within PA elements. For this purpose, this paper is structured as follows: Section 1 contains an introduction. Section 2 describes fundamentals and definitions. Section 3 presents existing approaches and assesses their deficits. Section 4 Results then presents the methodology and validates it. Section 5 concludes with a summary and outlook.

2. Fundamentals

To facilitate understanding of the following methodology, this section briefly describes product architecture, upgrade circular economy, as well as foresight and CV.

2.1 Product architecture (PA)

The PA consists of the product structure (PS) and the functional structure (FS), as well as the connecting transformation relationships [9,10]. The FS of a product represents the decomposition of the required main function into well-defined sub-functions. The solution-neutral formulation of the functions is essential for systematically allocating different solution variants within the PA. The PS, also known as the construction structure, models the physical composition of the product. It comprises assemblies that can be further subdivided into parts or components [10]. The design of the PA is a key determinant of a product's modifiability. Modular architectures enhance changeability by localizing modification within well-defined modules and, based on clearly specified interfaces, limiting the cross-module impact of changes on other modules and their functions [11]. Flexible, adaptive PAs are critical to commercial success because they robustly integrate dynamic customer and market requirements as well as late changes [6].

2.2 Upgrade Circular Economy (UCE)

By implementing sustainable value creation, the circular economy (CE) is an alternative to the take-make-use-waste principle of value creation [12]. CE is an economic system that aims to keep resources in circulation for as long as possible [13] while decoupling economic growth from resource consumption [12] and minimizing waste. As CE strategy, industrial remanufacturing reprocesses old products into as-new or higher-valued products by reassembling new, reconditioned, or updated parts and components after prior disassembly [14]. In contrast, the UCE extends these principles by establishing the circulation of products and components while preserving and potentially increasing their embedded value. The functional enhancement (FE) of the product through updates and upgrades ensures that the reprocessed product is once again competitive with new products [15], implying a higher innovation level than in the CE. One strategy for implementing this FE is re-assembly [14]. In UCE, the re-assembly must invariably yield upgraded value [3]. Thus, remanufacturing (CE) restores, re-assembly (UCE) upgrades and recombines modules.

2.3 Foresight and change variance

The increasing complexity and dynamism of the market environment necessitate the systematic inclusion of anticipated future requirements in product development. Uncertainty arises from evolving market and customer demands as well as ongoing technological advances. Strategic foresight is an appropriate method for anticipating such requirements [7]. It enables a comprehensive analysis of trends, technologies and influencing factors, thereby providing the basis for integrating evolving requirements into product design [8]. Operationally, this involves identifying modules that implement requirements expected to change and designing them for upgradability [7]. The concept of CV is defined as the temporal frequency of modifications to a product or its architecture. It is linked to the degree of change, which captures external

market dynamics and internal organizational perspective, and thereby characterizes the product context [2]. To assess predicted developments, the Technology Readiness Level (TRL) of NASA and the PESTLE analysis are further used. The TRLs are a systematic metric to assess the maturity of a technology [16]. In the PESTLE analysis, the environment is analyzed in six categories from a political, economic, sociocultural, technical, legal and ecological-geographical perspective, then rated on a scale [17]. We use the environmental analysis as the analysis of the technology's impact on the company.

3. Related Work

In advance of the development of the methodology existing literature was revised. Table 1 shows 13 approaches, which were identified through non-systematic database searches and snowballing, and their evaluation regarding seven evaluation criteria. Firstly, foresight methods must be employed to predict changes in PAs resulting from market and technological developments. Furthermore, the consideration of dynamics and CV is crucial for anticipating the frequency and nature of future alterations. The criteria also include the incorporation of sustainability as well as suitability for re-assembly, with relevance given to industrial feasibility, modularity and ability for disassembly of the product. Additionally, the approach should not only induce improvements in PAs but also be capable of integrating changes into it. Thus, identifying critical (PA) elements is essential. The approaches are evaluated according to predefined criteria using a Likert scale, which provides a discrete assessment of the degree to which a characteristic is present. Visualization with Harvey balls enables an direct qualitative assessment aligned with the Likert scale [18].

Table 1: Evaluation of related work (references of papers: [19–31])

Legend: ○ not fulfilled ◔ initially fulfilled ◑ partially fulfilled ◕ fulfilled ● fully fulfilled	Methods of foresight	Consideration of dynamics	Consideration of change variance	Consideration of sustainability	Suitability for re-assembly	Improvement of product architecture	Integration of improvements into product architecture	Identification of (critical) product architecture elements
DAGMAN & SÖDERBERG (2013)	○	○	○	●	◑	◕	◑	○
BAUER ET AL. (2014)	◕	◑	◑	○	○	○	○	◔
HOFFMANN (2017)	○	◑	◑	◔	◑	◕	◑	○
SANAEI ET AL. (2018)	○	○	○	○	◔	◕	◔	○
GREVE ET AL. (2019)	◕	◑	◑	○	◔	◑	◔	◔
YAMADA ET AL. (2020)	○	◔	○	◑	◑	◑	◔	○
MARTHALER (2021)	◕	◑	○	◑	◔	◔	◔	○
MOHSSENI (2021)	◕	◕	◑	◔	◔	◕	◕	○
DAHMANI ET AL. (2022)	○	○	○	●	◑	◔	○	○
RIESENER ET AL. (2023)	○	◔	○	◕	◑	◑	◔	○
GUO (2024)	◕	◑	◑	◑	◔	○	○	○
KUEBLER ET AL. (2024)	○	◑	◔	◔	○	◔	○	○
KUEBLER ET AL. (2024b)	◕	◔	◑	◑	◔	◑	○	◑

The following paragraph briefly outlines the approaches highlighted in Table 1. GREVE ET AL. introduce a method to identify new product features and assess their future customer value in modular product families

using conjoint analysis and Monte Carlo simulation, thereby deriving robust architectures [22]. MARTHALER outlines a five-step framework for future-oriented, cross-generational product development that combines long-term scenarios, as-is/to-be analyses and a cross-generational roadmap [27]. MOHSSENI develops a three-step approach for life-cycle-robust PAs that translates scenario-based customer needs through simplified QFD into functional requirements and uses DSM-based interdependence analysis and clustering to derive optimized architectures [28]. DAHMANI ET AL. present an integrative framework combining lean and eco design to support sustainable product development [21]. RIESENER ET AL. propose a four-step concept to modularize existing products for the CE and derive economically viable circular modules and modular PAs [29]. Perspectives are weighted through a pairwise comparison and consolidated into a heat map that exposes the modular architecture [29]. GUO proposes a four-step method to assess technology leap potential, developing a Technology Leap Radar, a Disruption Heat Map with entry probabilities and speeds, and using strategic as well as financial evaluation criteria [23]. KUEBLER ET AL. present a four-step roadmap approach whose final phase operationalizes changes in product properties and guides generational extensions by prioritizing future sub-functions and systems [26].

Across these approaches, sustainability is addressed most comprehensively by DAHMANI ET AL. and RIESENER ET AL., whereas foresight and CV are only selectively considered, mainly focusing on customer-side dynamics [22], technological disruption [23] or scenario-based long-term developments [27,28]. Systematic improvement of PAs and the integration of changes are central primarily in KUEBLER ET AL. [26] and MOHSSENI, while modular, upgrade-oriented design and thus suitability for re-assembly are supported mainly by RIESENER ET AL. and DAHMANI ET AL., and to a lesser extent, by the sources [22,23,26–28]. The explicit identification of critical elements is addressed explicitly only by GREVE ET AL. and KUEBLER ET AL. [26]. Summarizing, there is currently no approach that simultaneously considers sustainability, change variance, future market and technology developments, and systematic improvement of PAs, while also enabling the identification of critical PA elements. This motivates the research question of the paper: How can a methodology systematically predict the CV of PA elements to further identify critical PA elements?

4. Results

The aim of this paper is to develop a methodology that can identify the market- and technology-induced CV, to transfer market- and technology-induced CVs onto product architecture elements (PAEs) for determining the elements CV and to identify critical elements. As the methods for market and technology differ from each other, they are divided into two separate steps. Figure 1 shows the four-step concept of the methodology.

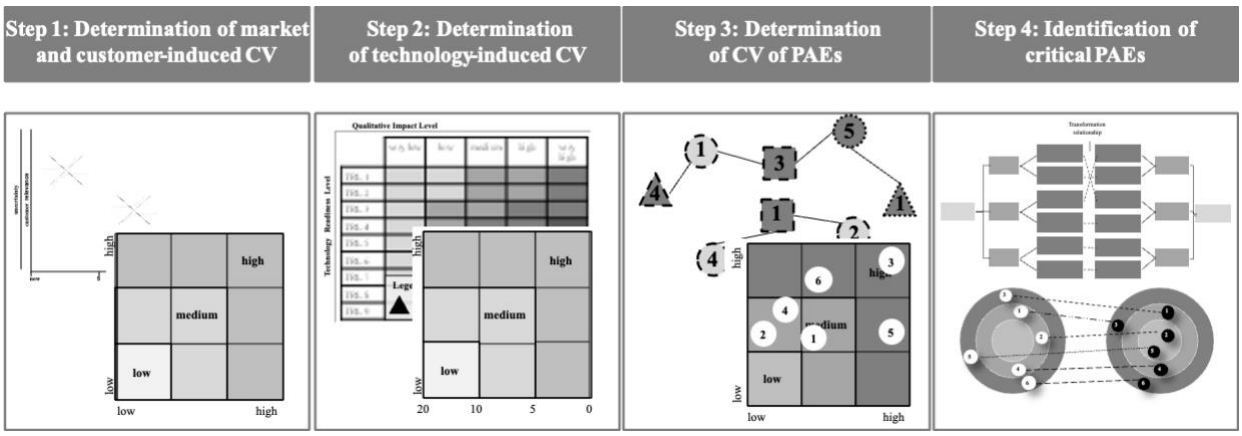


Figure 1: Four-step concept of the methodology

4.1 Determination of market and customer-induced CV

In the first step the market- and customer-induced CV is determined through the customer characteristic change potential (CCCP). The CCCP is determined in three sub-steps. First, the probability of change (POC)

of the customer characteristics (CCs) is defined through evaluation criteria. Dynamic market boundary constraints (for example competition) and additionally company-specific (see, e.g., [32,33]) and application-specific criteria (for example sustainability or resource consumption) (see, e.g., [34]) can be used as evaluation criteria. The QFD evaluation scale (0,1,3,9) is used to assess the influence of each evaluation criterion on each CC [35]. For each CC, the scores assigned across all evaluation criteria are summed and averaged to obtain a mean value that quantifies their influence. Depending on this influence, the POC is classified. Second, the customer relevance (CR) of the different CCs is determined using the Kano model. The CCs are classified into the Kano model categories basic, performance and delighter. Unfortunately, the Kano model cannot depict future evolutions. Therefore, an extension of the Kano model was developed, which shows evolution over time and CR in Appendix 1. In the third sub-step, the results of the first and second sub-step are transferred into a matrix to determine the extent of the CCCP into low, medium or high [36].

To validate the methodology, an exemplary application was carried out in each step. The PA was greatly simplified to focus on explaining the method. Market trends and technological trends are also required to apply the method and should be identified through adequate existing market and technology analyses. In the first step, the CCs are evaluated as shown in Table 2 according to a similar method by SCHIFFER [37]. Through a sum and mean value calculation the POC is determined. To determine CR, CCs were classified in the Kano model using examples, whereby the time of relevance and the characteristic value (basic, performance, delighter) had to be determined in advance using forecasting methods. CR can be quantified using Appendix 2 and calculated using the formula in Appendix 3. Table 3 shows the results of the CR calculation. PHAAL ET AL.'s time intervals are used, merging 10-15 and 15-20 years [38]. Industry specific calibration was omitted and may alter time intervals. The results of the POC and the CR of the CCs can be visualized in a matrix, which is shown in Figure 5, to determine the CCCP.

Table 2: Evaluation of POC of customer characteristics

Evaluation criteria	**Evaluation** 0 – no influence 1 – low 3 – medium 9 – high	**Customer characteristic(CC)**			
		CC 1	CC 2	CC 3	CC 4
	Dynamic	0	9	3	3
	Degree of influence	3	3	1	0
	Demand	1	9	0	3
	Competition	9	3	1	9
	Total assets	13	24	5	15
	Mean value	3,25	6	1,25	3,75
	Probability of change	medium	high	low	medium

Table 3: Evaluation of customer relevance

Customer characteristic(CC)	Characteristicvalue in the Kano model	Kano model weighting	Time of relevance of the feature	Time factor	Corporate strategy	Calculated relevance	Resulting Customer relevance
CC1	Performance	2	0-5 years	2.0	-	4	medium
CC2	Performance	2	10-20 years	1.0	-	2	low
CC3	Delighter	1	0-5 years	2.0	+0.5	2.5	medium
CC4	Basic	3	5-10 years	1.5	-	4.5	high

4.2 Determination of technology-induced CV

In the second step, the determination of the technology-induced CV takes place. It is divided into the three sub-steps determination of the probability of change of the technology, the entry point (EP) of the technology and its disruption potential (DP). The POC of the technology is identified with its TRL [16] and its PESTLE analysis [17]. In the PESTLE analysis, technologies are classified into six categories and then quantified using a specially developed calculation [17]. The results of the PESTLE analysis and the TRL are transferred to a detailed matrix, which then classifies the POC as low, medium or high. The TRL is also used to determine the EP of a technology into the market. It is assumed that the higher the barriers to entry, the later a technology will enter the market. Accordingly, time stages are linked to the TRL to make an initial

assessment of the EP. Finally, the DP is determined in a matrix with the POC of the technology on the ordinate and the EP on the abscissa.

In the example, the POC of technologies is determined through TRL and PESTLE analysis. The TRL are classified according to NASA's nine-point scale [16] and could be visualized in a technology radar, which is shown in Figure 2 for technology examples. In the PESTLE analysis, the technologies are considered in the six categories and rated on a positive or negative QFD scale as low (1/-1), medium (3/-3), and high (9/-9) . The numerical values are used for quantification, whereby the impact is defined as the sum of the ratings for the six categories. The sum is then classified on a scale, thus assigning a qualitative impact level (QIL) to each technology. The QIL can be very low (less than or equal to 0), low (1-6) (max. 6*1), medium (7-18) (max. 6*3), high (19-36) (max. 6*6) or very high (37-54) (max. 6*9). These results can be used to determine the POC of the technology being adopted according to the scale shown in Figure 3. The EP of the technology is determined using the scale in Appendix 4, with high TRL indicating earlier entry into the market and low TRL indicating later entry. A lower TRL generally implies higher barriers to adoption and therefore a longer time to widespread market uptake [23]. The CV is low when modifications to the PA occur infrequently. Consequently, low QIL and low TRL are associated with low frequency of change. The results of the POC and the EP of the technology are used to determine the DP in the matrix in Figure 5.

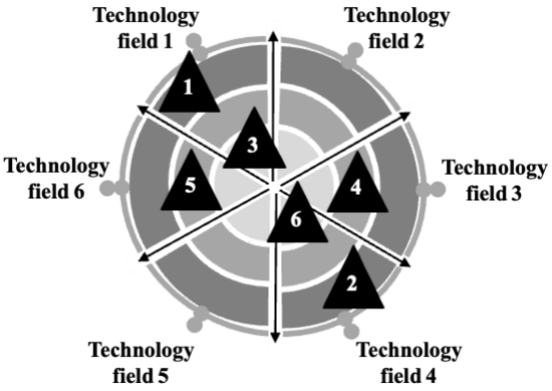


Figure 2: Exemplary technology radar

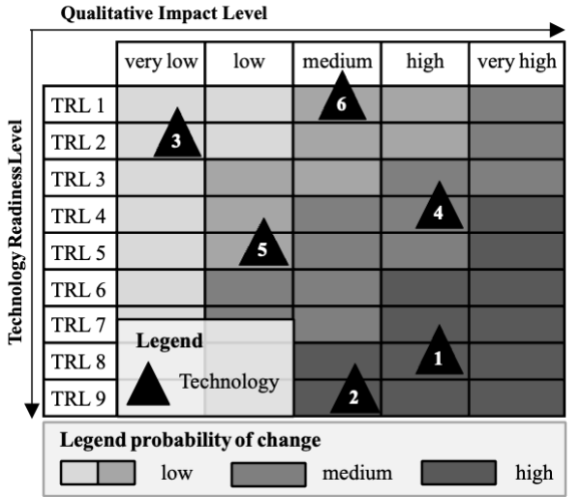


Figure 3: Determination of POC of technologies

4.3 Determination of CV of PAEs

The results of the first and the second step are used in the third step to determine the CV of PAEs. This involves two sub-steps. First, the customer and technological characteristics must be linked to the FS of the product. For this purpose, the FS of the product must be known and represented at the configurable level. Subsequently, the link between features and functions can be visualized. Secondly, the CVs of the product functions are classified and then visualized. Therefore, a resulting CV is first formed out of the CVs of the characteristics associated with the function using the scale in Appendix 5. If there are more than two characteristics, a mean value calculation is used to determine the result. The resulting CV can then be entered into the connection diagram of the first sub-step to complete the visualization. The CV results of the functions are then presented in the change variance matrix (CVM), which compares the CCCP and the DP. Using this representation, the results can be easily transferred to the shell model of the market perspective. The shell model serves to harmonize product and market perspectives [2]. The market perspective examines the functional side of the product in this case and thus determines its flexibility [2].

In the example, the market and technological characteristics are linked to the functional structure. Since the entire product architecture is required in addition to the functional structure, the exemplary product architecture is visualized in Appendix 6. The characteristics are visually linked to the functions they

influence with adequate existing market and technology analyses. The resulting CV in the function is determined through these links. According to the scale in Appendix 5, if there are two characteristics, the resulting CV will always be the highest of the two characteristics. For more than two characteristics, a mean value calculation is used. With this information, the CV of the functions is determined in tabular form (Table 4) and then visualized in the connection diagram (Figure 4).

Table 4: Determination of function's CV

Function	Number of technology	Change variance of technology	Number of customer characteristic	Change variance of customer characteristic	Change variance of function
1	4	medium	3	medium	medium
2	6	low	1	medium	medium
3	2	high	4	high	high
4	3	low	1	medium	medium
5	1	high	3	medium	high
6	5	medium	2	high	high

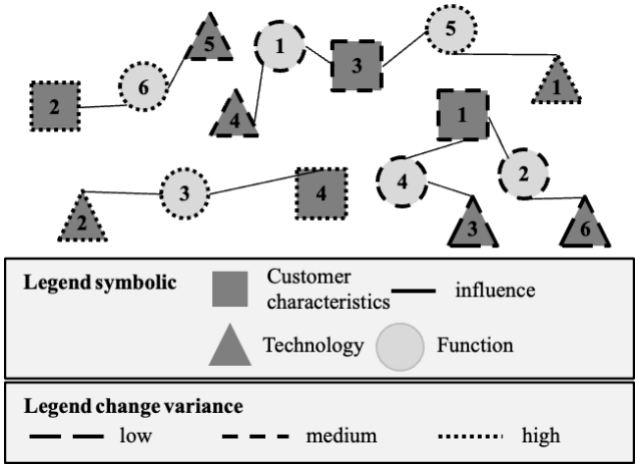


Figure 4: Connection diagram with change variance of functions

The results of the first and second step, the CCCP and the DP of the technology, are used in the CVM. The shell model is created out of the CVM and shows the necessary flexibility of the functions in the product life cycle. Figure 5 shows the transfer of the CCCP and the DP results to the CVM and the shell model.

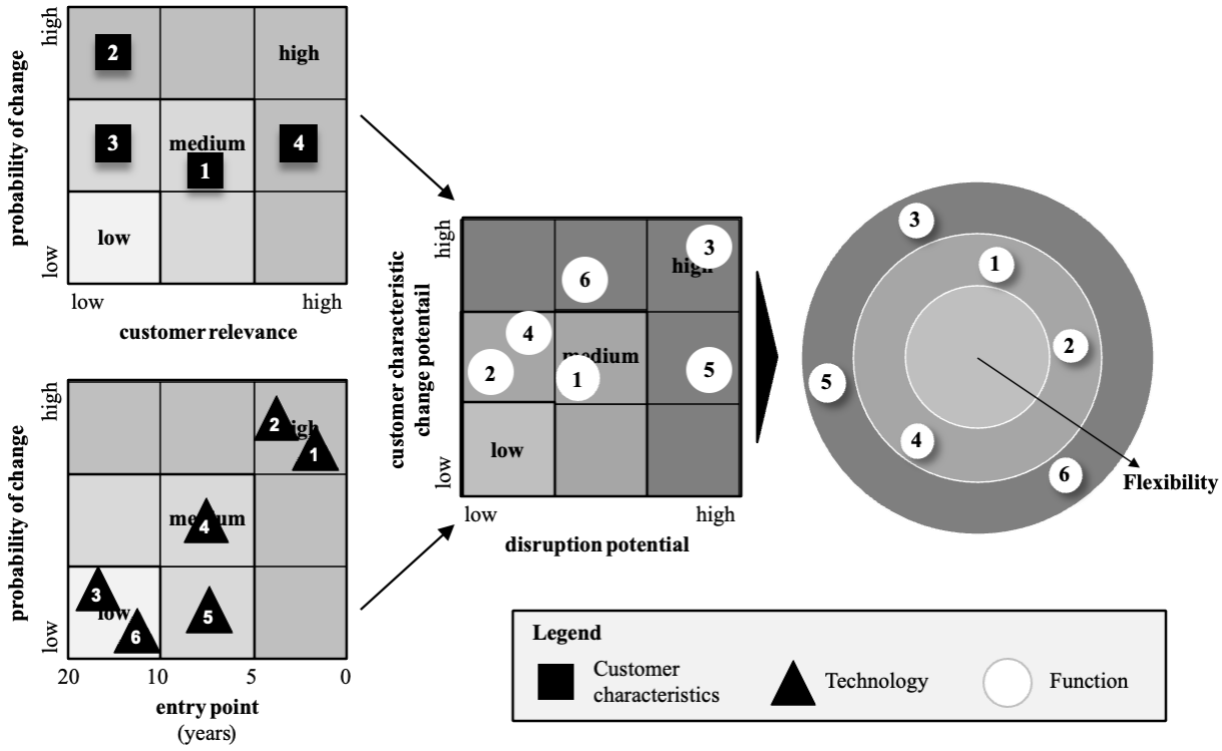


Figure 5: Transfer of CCCP and DP into the CVM and shell model

4.4 Identification of critical PAEs

In the final step, the critical PAEs are identified. The first sub-step involves determining the flexibility of the PS, which requires first defining and presenting the PS. The shell model of the product perspective is then used to determine the flexibility of the PS elements [2]. The elements are positioned in the shell model

using qualitative criteria, operationalized through a morphological decision matrix that evaluates product-related components according to multiple criteria [2]. The second sub-step determines the discrepancies in the flexibility of PS (components) and FS (functions) to derive the criticality of the PS elements. First, the transformation relationships between PSs and FSs must be defined. Then it can be determined whether there is a discrepancy in the flexibility of connected elements. Without a discrepancy, the element is not critical. With a discrepancy in flexibility, criticality must be determined based on the following explanations. If the flexibility of the component is higher than the flexibility of the function, the PAE is not critical and there is ineffective flexibility. PAEs are (very) critical, if the component is less flexible than the function requires. These results of the determination of critical PAEs can be presented in a comparison of the shell models, as shown in Figure 6.

For determining the critical PAEs in the example, the PS in Appendix 6 must be classified in the shell model according to the flexibility available for the components. FS and PS in the shell model are then linked with the transformation relationships to determine discrepancies in flexibility. This creates the results shown in Figure 6 for the criticality of the PAEs. As an example, function 5 and component 5 are very critical, which shows that future developments will need changes in component 5 to fulfill upcoming trends. The same applies to the other critical elements. For an example based on a car, see Appendix 7 and Appendix 8.

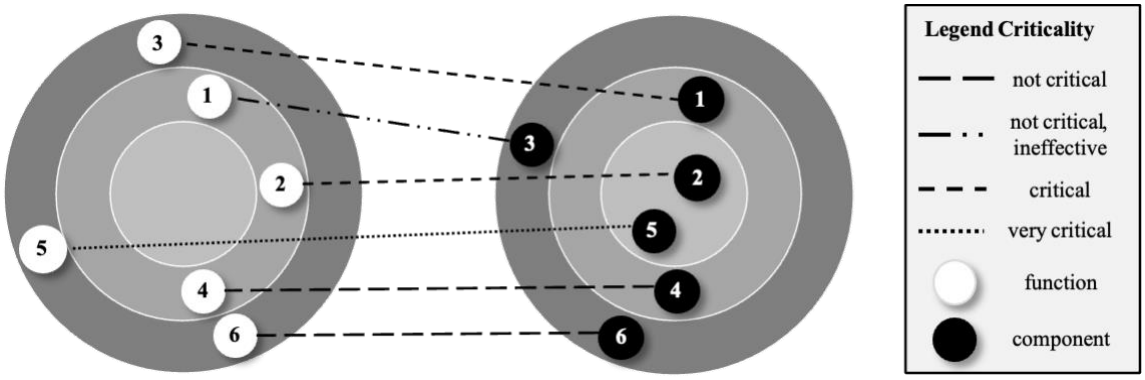


Figure 6: Criticality of product architecture elements

In summary, Section 4 identified the CV of CC and technologies, linked these to the FS of the PA and identified the critical elements of the PA using the shell model of the PS. This enables an assessment of future changes in PAs and challenges associated with the PA.

5. Summary and Outlook

This paper proposes a methodology to determine change variance in product architecture elements. A four-step approach is proposed and applied to a generic PA. The results show that the methodology can be applied onto PAs and is able to visualise critical PAE's. Each step of the methodology was described and illustrated with an example. The market-induced and technology-induced CV is determined through foresight-based analyses and calculations for customer characteristics and technologies. These results are used in the third step to determine the CV of the functional PS. The last methodology step locates elements in functional and structural PAs that are most susceptible to change and therefore critical to the product's longevity.

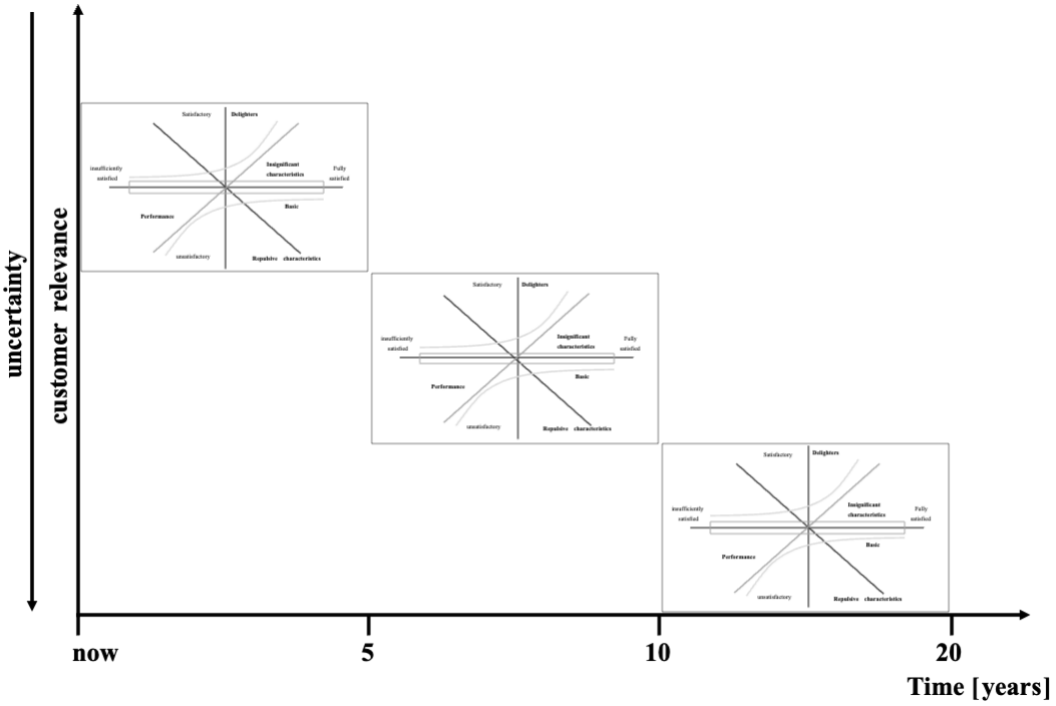
Manufacturing companies could use the approach for proactive integration of foreseen changes into their PAs in early product development phases. It enhances upgradability, extends the product lifetime and supports the practical realization of (U)CE principles. With the results of this methodology, it becomes possible to systematically change PAs and optimize them for future developments. In future research, the extent to which the methodology is suitable for more complex and large-scale PAs should be investigated. Moreover, its applicability in industrial practice could be evaluated, for example through empirical case studies or pilot implementations.

Acknowledgements

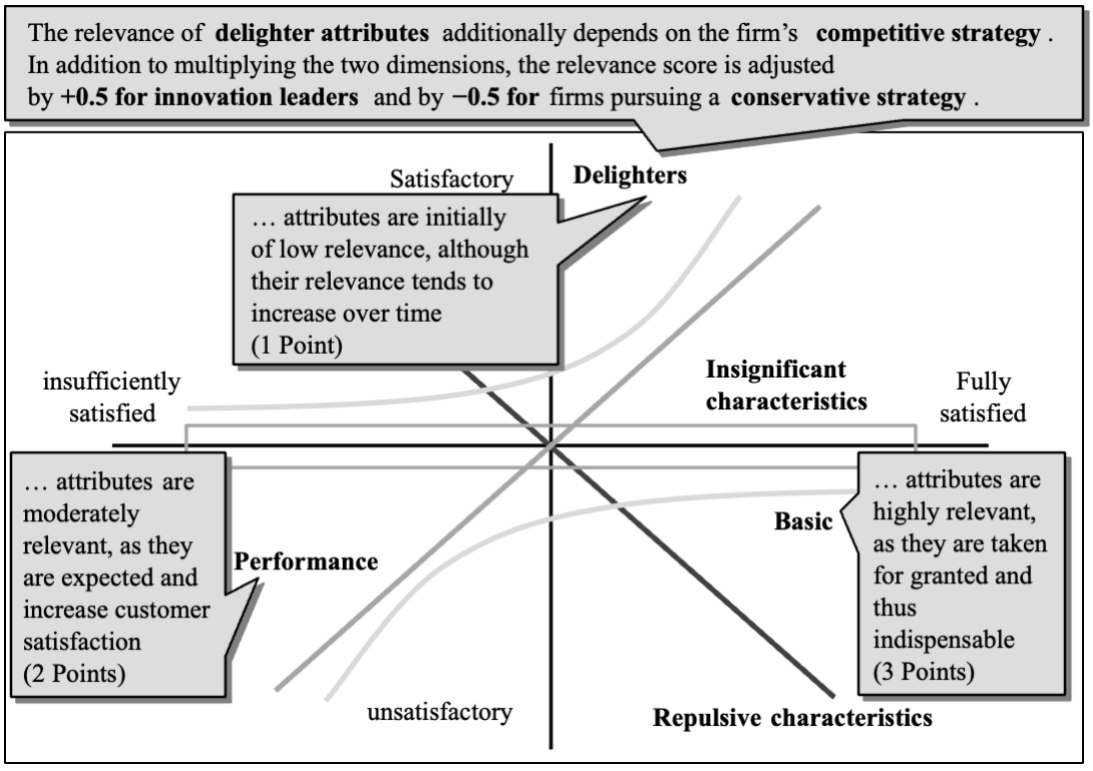
Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

Appendix

Appendix 1: Extended Kano model



Appendix 2: Kano model with explanations for quantification



Appendix 3: Calculation of customer relevance

$$CuR_a = KB_a * TF_a + CS_{a_{de}}$$

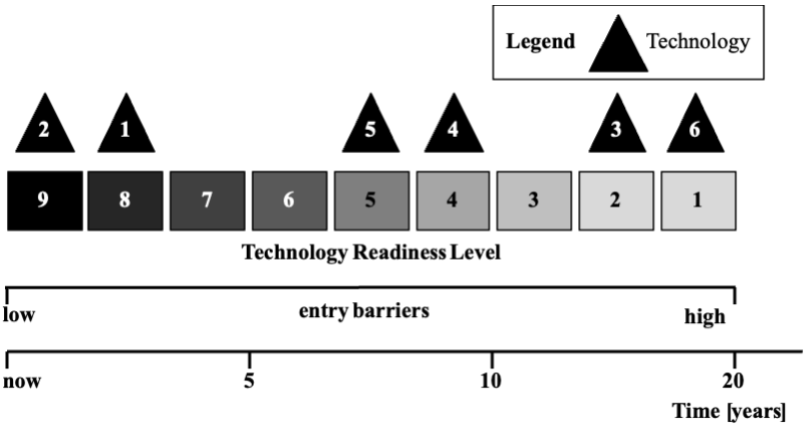
CuR_a = numerical customer relevance of customer characteristic a

KB_a = Kano Basic Factor of customer characteristic a (basic = 3 performance = 2 delighter = 1)

TF_a = Time factor of customer characteristic a (0-5 years = 2 5-10 years = 1.5 10-20 years = 1)

$CS_{a_{de}}$ = Evaluation of company's strategy of delighter characteristic a_{de}
(innovators = +0.5 conservative = -0.5)

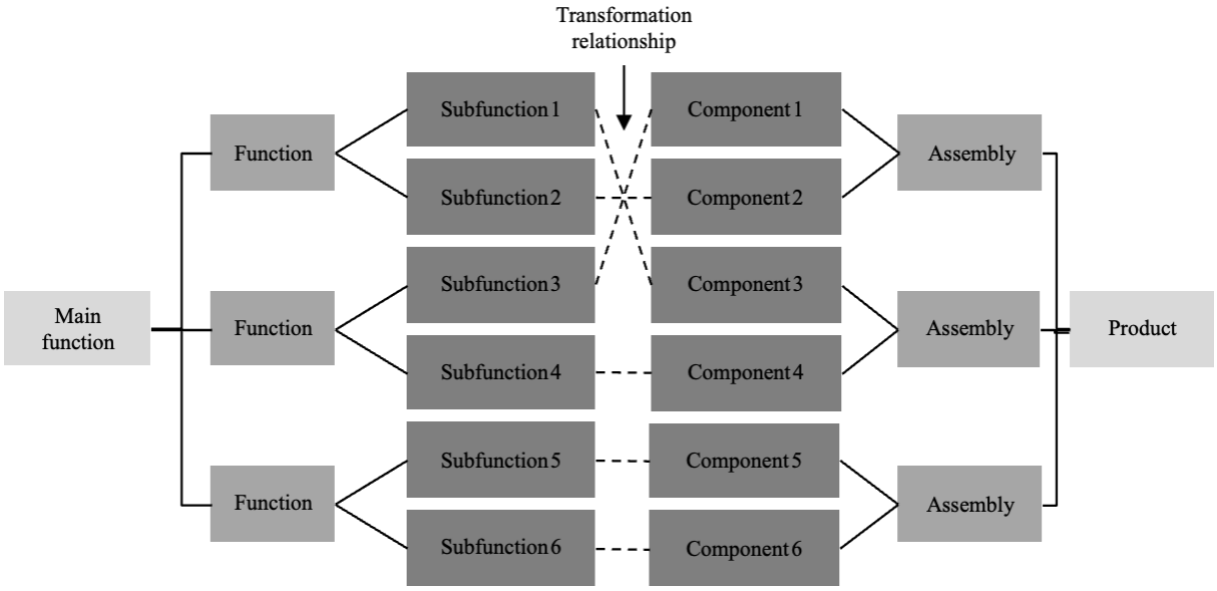
Appendix 4: Scale for entry point of technology



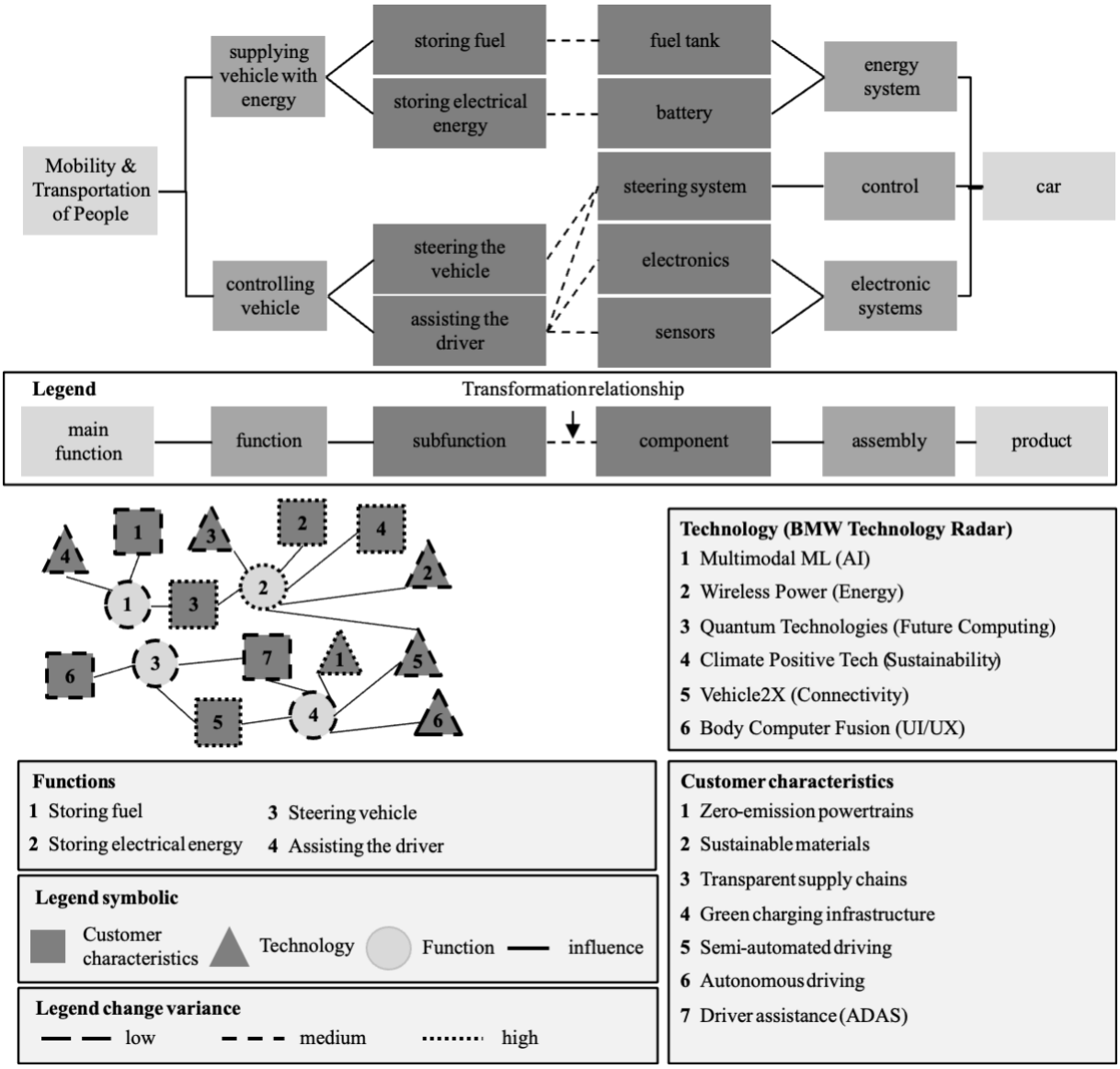
Appendix 5: Scale for resulting change variance of the functional structure

Customer characteristic change potential (CCCP) (change variance of customer characteristic)	Disruptions potential (DP) (change variance of technology)	Resulting change variance of function
low	low	low
low	medium	medium
low	high	high
medium	low	medium
medium	medium	medium
medium	high	high
high	low	high
high	medium	high
high	high	high

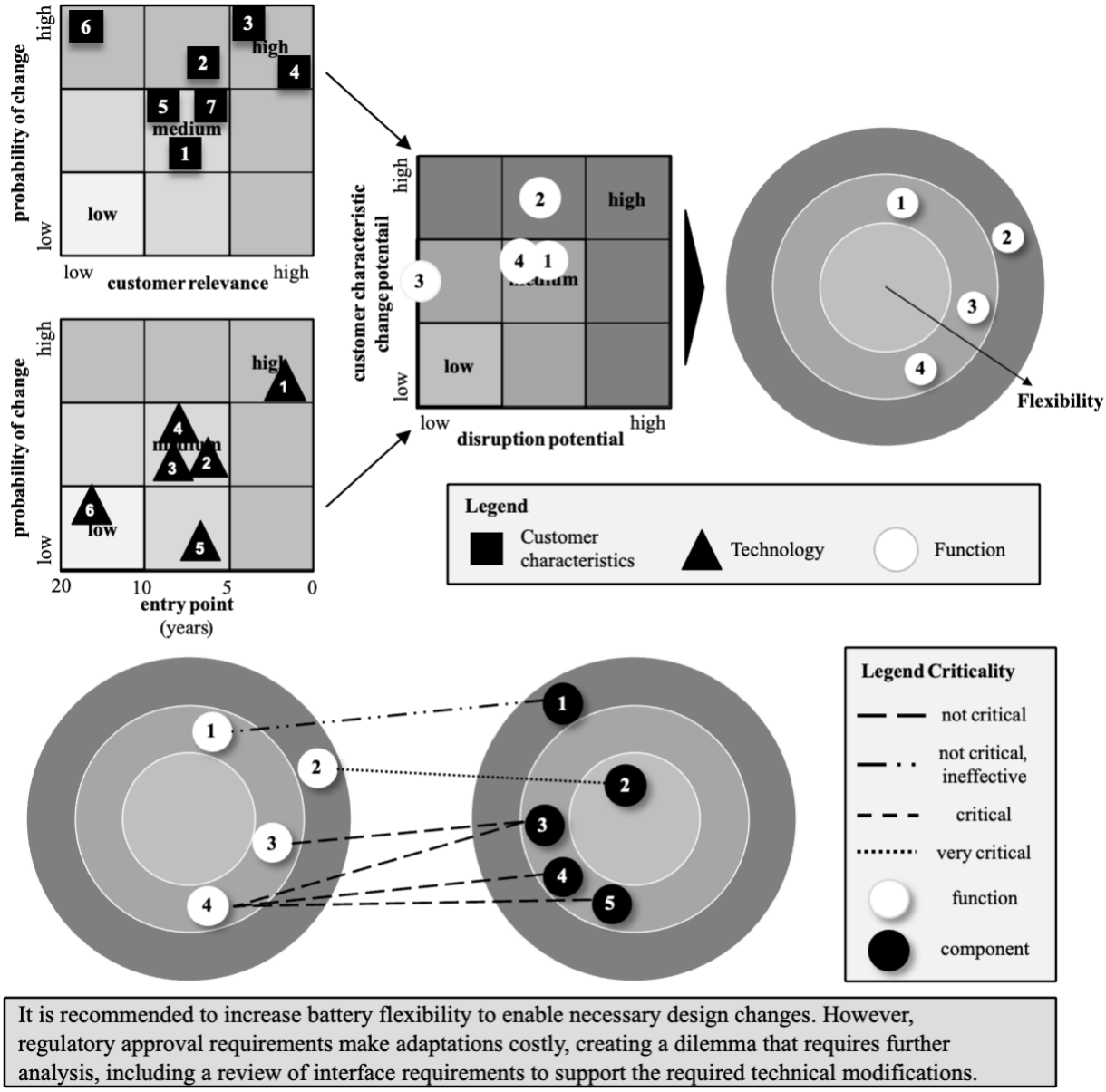
Appendix 6: Exemplary product architecture



Appendix 7: Example passenger car – reduced PA and connection diagram [39]



Appendix 8: Example passenger car – Transfer of CCCP and DP into the CVM and shell model & criticality of PAEs



It is recommended to increase battery flexibility to enable necessary design changes. However, regulatory approval requirements make adaptations costly, creating a dilemma that requires further analysis, including a review of interface requirements to support the required technical modifications.

References

[1] Riesener, M., Kuhn, M., Ruschitzka, C., Becker, D., Schuh, G., 2022. Identification of Fields of Action for the Development of Ecologically Sustainable Products, in: Technology management and leadership in digital transformation - looking ahead to post-COVID era. PICMET '22 : Portland International Conference on Management of Engineering and Technology : proceedings. 2022 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA. 8/7/2022 - 8/11/2022. IEEE, Piscataway, NJ, pp. 1–10.

[2] Schuh, G., Dölle, C., 2021. Sustainable Innovation: Nachhaltig Werte schaffen, 2. Auflage ed. Springer Vieweg, Berlin, Heidelberg, 306 pp.

[3] Riesener, M., Keuper, A., Schulz-Isenbeck, T., Schuh, G., 2025. Concept for Product Architecture Evolution Planning in the Context of Circular Economy. Procedia CIRP 136, 38–43.

[4] Hermann, A., Riesener, M., Schmitz, S., Neumann, H., 2025. Cost Model for Evaluation of Industrial Re-Assembly Processes, in: Kohl, H., Seliger, G., Dietrich, F., Vien, H.T. (Eds.), Decarbonizing Value Chains. Proceedings of the 20th Global Conference on Sustainable Manufacturing (GCSM 2024), October 9–11, 2024,

Ho Chi Minh City, Vietnam, 1st ed. 2025 ed. Springer Nature Switzerland; Imprint Springer, Cham, pp. 563–571.

[5] Borgmann, F., Henke, M., 2016. Modulare Produktion, in: Spath, D., Westkämper, E. (Eds.), Handbuch Unternehmensorganisation. Strategien, Planung, Umsetzung, Living reference work, continuously updated edition ed. Springer Vieweg, Berlin, Heidelberg, pp. 1–22.

[6] Fricke, E., Schulz, A.P., 2005. Design for changeability (DfC): Principles to enable changes in systems throughout their entire lifecycle. Syst. Engin. 8 (4).

[7] Kuebler, M., Schuster, W., Schwarz, S.E., Braumandl, A., Siebe, A., Albers, A., 2023. Upgradeable Mechatronic Systems - An Approach to determine changing Product Properties using Foresight. Procedia CIRP 119, 78–83.

[8] Mörtl, M., 2003. "DESIGN FOR UPGRADING" OF MACHINES AND PRODUCTION PROCESSES: A GUIDLINE BASED ON ACTUAL DEMANDS OF INDUSTRY AND SUSTAINABLE DESIGN, in: INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN. ICED, Stockholm. 19-21 August, pp. 1–10.

[9] Schuh, G., Lenders, M., Nußbaum, C., 2012. Produktarchitekturgestaltung, in: Schuh, G. (Ed.), Innovationsmanagement, 2., vollst. neu bearb. und erw. Aufl. ed. Springer Vieweg, Berlin, pp. 115–160.

[10] Schuh, G., Riesener, M., 2018. Produktkomplexität managen: Strategien - Methoden - Tools, 3., vollständig überarbeitete Auflage ed. Hanser, München, 273 pp.

[11] Ulrich, K.T., Eppinger, S.D., Yang, M.C., 2020. Product design and development, Seventh edition, International student edition ed. McGraw-Hill, New York, NY, 432 pp.

[12] Schlund, S., 2016. Nachhaltige Wertschöpfung, in: Spath, D., Westkämper, E. (Eds.), Handbuch Unternehmensorganisation. Strategien, Planung, Umsetzung, Living reference work, continuously updated edition ed. Springer Vieweg, Berlin, Heidelberg, pp. 1–12.

[13] Genders, S., 2025. Nachhaltigkeit: Grundlagen und Bedeutung, in: Genders, S. (Ed.), Nachhaltig gründen. Ein Praxisleitfaden für verantwortungsvolle Unternehmensgründungen. Springer Gabler, Berlin, Heidelberg, pp. 3–18.

[14] Schuh, G., Schmitz, S., Schopen, M., Neumann, H., 2023. Concept For Databased Sales And Resource Planning For Re-Assembly In The Automotive Industry. Conference on Production Systems and Logistics, 554–562.

[15] Schuh, G., Schmitz, S., Lukas, G., Niwar, L., Welsing, M., Calchera, R., 2023. Ordnungsrahmen für eine zirkuläre Produktionswirtschaft.

[16] Mankins, J.C., 1995. Technology Readiness Level. NASA. http://www.artemisinnovation.com/images/TRL_White_Paper_2004-Edited.pdf. Accessed 16.12.25.

[17] Kaufmann, T., 2025. Strategiewerkzeuge aus der Praxis: Band 1: Analyse und Beurteilung der strategischen Ausgangslage, 2., aktualisierte Auflage ed. Springer Gabler, Berlin, 303 pp.

[18] Likert, R., 1932. A technique for the measurement of attitudes, 22nd ed., New York.

[19] Bauer, W., Werner, C., Elezi, F., Maurer, M., 2014. Forecasting of Future Developments Based on Historic Analysis, in: , DS 77: Proceedings of the DESIGN 2014 13th International Design Conference, pp. 735–748.

[20] Dagman, A., Söderberg, R., 2013. Toward a Method for Improving Product Architecture Solutions by Integrating Designs for Assembly, Disassembly and Maintenance, in: Proceedings of the ASME International Mechanical Engineering Congress and Exposition 2012. Presented at ASME 2012 International Mechanical Engineering Congress and Exposition; November 9 - 15, 2012, Houston, Texas, USA. ASME 2012 International Mechanical Engineering Congress and Exposition, Houston, Texas, USA. 11/9/2012 - 11/15/2012. ASME, New York, NY, pp. 377–387.

[21] Dahmani, N., Belhadi, A., Benhida, K., Elfezazi, S., Touriki, F.E., Azougagh, Y., 2022. Integrating lean design and eco-design to improve product design: From literature review to an operational framework. Energy & Environment 33 (1), 189–219.

[22] Greve, E., Rennpferdt, C., Hartwich, T., Krause, D., 2019. Determination of Future Robust Product Features for Modular Product Family Design, in: Volume 14: Design, Systems, and Complexity. ASME 2019 International Mechanical Engineering Congress and Exposition, Salt Lake City, Utah, USA. 11.11.2019 - 14.11.2019. American Society of Mechanical Engineers.

[23] Guo, D., 2024. Methodik zur Potenzialbewertung von Technologiesprüngen. Dissertation, Aachen.

[24] Hoffmann, C.-A. (Ed.), 2017. Methodik zur Steuerung modularer Produktbaukästen. Dissertation. Springer Fachmedien Wiesbaden, Wiesbaden.

[25] Kuebler, M., Lange, C., Glasmacher, B., Düser, T., Albers, A., 2024. Future-robust product design – validating influencing factors on upgradeable mechatronic systems. Proc. Des. Soc. 4, 633–642.

[26] Kuebler, M., Taghavian, K., Teufel, F., Siebe, A., Albers, A., 2024. Upgradable Mechatronic Systems - Roadmapping for changing Product Properties, in: , ISPIM Innovation Conference, Tallinn, pp. 1–13.

[27] Marthaler, F., 2021. Zukunftsorientierte Produktentwicklung - Eine Systematik zur Ableitung von generationsübergreifenden Zielsystemen zukünftiger Produktgenerationen durch strategische Vorausschau. Dissertation, Karlsruhe.

[28] Mohsseni, A., 2021. Gestaltung von lebenszyklusrobusten Produktarchitekturen am Beispiel des Lithium-Ionen-Batteriesystems. Dissertation, 1. Auflage ed.

[29] Riesener, M., Kuhn, M., Keuper, A., Hellwig, F., Schuh, G., 2023. Concept for Product Modularization in the Context of Circular Economy, in: 2023 Portland International Conference on Management of Engineering and Technology (PICMET). 2023 Portland International Conference on Management of Engineering and Technology (PICMET), Monterrey, Mexico. 23.07.2023 - 27.07.2023. IEEE, pp. 1–9.

[30] Sanaei, R., Otto, K., Wood, K., Hölttä-Otto, K., 2018. A rapid algorithm for multi-objective Pareto optimization of modular architecture, in: Design methods and tools. DS 87-4 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 4: Design Methods and Tools, Vancouver. 21.-25.08.17. Curran Associates Inc, Red Hook, NY, pp. 169–178.

[31] Yamada, S., Miyajima, S., Yamada, T., Bracke, S., Inoue, M., 2020. Decision Support Method for Upgrade Cycle Planning and Product Architecture Design of an Upgradable Product. IJAT 14 (6), 919–929.

[32] Homburg, C., 2020. Marketingmanagement: Strategie - Instrumente - Umsetzung - Unternehmensführung, 7., überarbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden, Heidelberg, 1473 pp.

[33] Kotler, P., Keller, K.L., Chernev, A., Opresnik, M.O., 2023. Marketing-Management: Konzepte - Instrumente - Unternehmensfallstudien, 16., aktualisierte Auflage ed. Pearson, München, 666 pp.

[34] Belton, V., Stewart, T.J., 2002. Multiple Criteria Decision Analysis. Springer US, Boston, MA.

[35] International Standard, 2021. Application of statistical and related methods to new technology and product development process: Part 1: General principles and perspectives of quality function deployment (QFD).

[36] Stevens, S.S., 1946. On the Theory of Scales of Measurement. Science 103 (2684), 677–680.

[37] Schiffer, M.G., 2013. Szenariorobuste Produktarchitekturgestaltung. Zugl.: Aachen, Techn. Hochsch., Diss, 2013, 1. Aufl. ed. Apprimus-Verl., Aachen, 257 pp.

[38] Phaal, R., Farrukh, C.J., Probert, D.R., 2004. Technology roadmapping—A planning framework for evolution and revolution. Technological Forecasting and Social Change 71 (1-2), 5–26.

[39] BMW Group, 2025. Technologie Trend Radar. https://www.bmwgroup.com/de/innovation/technologie-trend-radar.html. Accessed 22 September 2025.

Biography

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh (*1958) has the Chair of Production Engineering at RWTH Aachen University and is a member of the Board of Directors of the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen and the Fraunhofer Institute for Production Technology IPT in Aachen. As a scientist and entrepreneur, Prof. Günther Schuh focuses on disruptive innovations, information and production management, and sustainable mobility solutions.

Dr.-Ing. Alexander Keuper (*1991) studied mechanical engineering at RWTH Aachen University and has been a research assistant at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen since 2019, where he received his PhD and was appointed head of the Innovation Management Department in 2024. Dr.-Ing. Alexander Keuper works at the intersection of complexity management, development management, and data intelligence to support sustainable innovation under increasing system complexity.

Frederike Hellwig (*1997) studied business administration and mechanical engineering at RWTH Aachen University and has been a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2022, where she has been working as a group lead development management since 2025. Frederike Hellwig focuses on the development of sustainable and circular product architectures and on the management of product complexity in industrial contexts.

Kira Rimbach (*1998) studied mechanical engineering at RWTH Aachen University and has been a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2025. Kira Rimbach focuses on operationalizing sustainability in industrial contexts and on developing approaches to address the associated complexity of product and systems.

8th Conference on Production Systems and Logistics

Concept For Designing Upgradeable Product Architectures In The Context of Upgrade Circular Economy

Frederike Hellwig[1], Alexander Keuper[1], Günther Schuh[1,2]

[1] *Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Aachen, Germany*
[2]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

The predominant linear economy model of manufacturing, characterized by a take–make–dispose logic, is increasingly incompatible with sustainability goals and resource limitations. Transitioning towards a circular economy requires a fundamental redesign of products to extend their functional lifespan and retain value over multiple use cycles. One key enabler is the capability of product architectures to integrate upgrades that improve performance or functionality. However, most existing architectures are not systematically designed for adaptability, which limits their circular potential. This paper introduces a framework for designing upgradeable product architectures in the context of upgrade circular economy. A structured five-step methodology supports early design decisions by (1) synthesizing design strategies such as modularization, interface standardization, and functional decoupling; (2) the assessment of interface flexibility; (3) a forecast of external change variance based on changing market demands and technological disruption; (4) identifying critical components and (5) iteratively applying targeted design strategies to improve upgrade circularity. A quantitative upgradeability score enables transparent comparison of architecture alternatives. The concept follows an iterative, model-based process integrating technical, market-based, and ecological perspectives. It provides a transferable framework for planning upgrades, supporting modular reassembly, and extending product lifecycles, thereby contributing to the practical implementation of upgrade circular economy principles in industrial product development.

Keywords

Upgrade Circular Economy; Circular Economy; Sustainability; Product Architecture; Product Design

1. Introduction

Climate change and the increasing pressure on natural resources make it clear that economies and societies need to significantly reduce their resource use and organise it in a more circular manner [1]. According to the *"Circularity Gap Report 2025"* [2], the global circularity rate has decreased to 6.9%, down from 7.2% in the previous year. At the same time, only 11.2% of all materials entering the economy are recycled, which highlights substantial inefficiencies in current waste management and recycling systems. Taken together, these figures suggest that the world economy still operates predominantly in a linear mode, with implications for resource use, environmental impacts, supply security and the need for more ambitious circular economy strategies. Against this background, recent EU legislation on net-zero manufacturing links industrial decarbonisation with improved product sustainability. The regulation promotes the consideration of criteria such as durability, ease of repair and maintenance and upgradability. In this context, the circular economy has emerged as a key concept for shifting from traditional linear economic models towards more sustainable forms of value creation [3]. Unlike the linear economy, which typically ends with product disposal, the

DOI: https://doi.org/10.15488/21008
ISSN: 2701-6277

circular economy emphasises the recovery and regeneration of products and materials and shifts the focus away from end of life scenarios [4]. In particular, circular strategies support more efficient resource use and longer product lifetimes [5]. Manufacturers must ensure that products remain viable in the market over extended periods without compromising economic or ecological goals [6]. This is particularly challenging in view of shorter product lifecycles driven by ongoing technological advances and rising customer expectations [7]. One way to address this tension is the implementation of value adding upgrades, which extend product lifetimes, postpone the need for entirely new products and foster an upgrade oriented circular economy [5]. Product architectures play a central role in enabling such upgradability, as they determine how easily functions and technologies can be adapted or extended over time, yet there is still no widely accepted design framework for upgradable products in a circular economy context [8]. Therefore, the overarching objective of this paper is to develop a framework for designing upgradeable product architectures in the context of the upgrade circular economy. The proposed methodology includes an approach that evaluates the flexibility of interfaces and also supports the prediction of market- and technology-driven change requirements in order to identify critical components. On this basis, targeted design strategies are applied iteratively to increase upgradeability. A quantitative upgradeability assessment enables a transparent comparison of different architectural alternatives. The framework is primarily intended for complex, technology-driven industrial products, where upgrades and evolving functions are relevant. The paper is structured as follows. After this introduction, Section 2 provides an overview of the terminology relevant to this work. Section 3 presents existing scientific approaches and derives the research gap. Based on the research gap, section 4 introduces the proposed framework for designing upgradeable product architectures.

2. Relevant Terminology

2.1 Upgrade Circular Economy

The circular economy is a theoretical concept that aims to decouple economic growth from resource consumption [9]. It represents an interdisciplinary reorientation of established resource and waste management strategies [10] and seeks to increase resource productivity by slowing, closing and reducing material loops [11]. Unlike the linear economy, it is restorative and regenerative in character [12] and places the entire product life cycle at the centre of consideration [13]. By keeping resources in closed loops, waste is minimised and the demand for primary raw materials decreases, as end-of-life products can serve as a source of materials for new ones [14].

Upgrade circular economy is an approach within the circular economy that not only aims to close material loops but explicitly seeks to increase product value through upgrades. Instead of merely preserving a product's original use or recycling it for additional life cycles, it focuses on extending product lifespan and functionality via updates and upgrades [15]. Upgrading is typically carried out on products that are still in good condition but whose capabilities no longer meet changed requirements, improving their functional and or cosmetic characteristics and restoring their competitiveness relative to new products [16]. By integrating new functionalities and technologies into existing products over time, this approach keeps products attractive and useful for customers and can increase customer satisfaction and loyalty [17]. Studies indicate that replacing around 10–30% of a product's components can already make it equivalent to a new product and achieve resource and emission savings of up to 50% [18]. The literature links the upgrade circular economy to sustainable manufacturing, increased profitability and innovation [19] and also associates it with new business models such as product-as-a-service [20]. Furthermore, several authors emphasise that its implementation requires an understanding of market dynamics, evolving customer needs and anticipated technological developments in order to design product architectures tailored to circular products [5].

2.2 Product Architecture and Modularization

Product architecture describes the elements of a product and their interrelationships for its design in terms of geometry, technology, and function [21]. It comprises a functional structure, a product structure, and the transformation relationships between them, representing both the functional and physical views of the product [22]. The functional structure decomposes the overall function into solution-neutral subfunctions, which are typically arranged hierarchically [23]. The product structure maps the physical composition into assemblies, components, and parts and supports the structuring of the product and the design process as well as the specification of component relationships [24]. When defined generically, it also enables the reuse of components or modules across product generations and product lines, thereby supporting circular economy strategies [25]. Transformation relationships link functional and physical structures and are characterized by the degree of functional and physical independence between subfunctions and components [21]. Different levels of functional and physical independence lead to different forms of modularity. Integral architectures exhibit strong interdependencies that complicate disassembly and upgrades [26], whereas modular architectures comprise distinct modules with clearly defined functions, which simplify disassembly, replacement and upgrading and thus support upgrade-oriented circular economy strategies [5]. Consequently, the design of the product architecture is essential for effectively extending product lifecycles through upgradability within the circular economy framework [27]. Product modularization, primarily used in variant management, structures products into modules according to defined criteria so that companies can manage increasing complexity from individualized customer requirements while maintaining low internal complexity [23]. The goal is an appropriate, rather than maximal, degree of modularization that balances standardization and individualization [28]. In a circular economy context, modularization facilitates the efficient implementation of market- and technology-driven upgrades [16].

3. Related Work

This section presents selected approaches from the literature that address circular economy strategies in connection with product architecture. The aim is to identify the research gap in the context of upgradable product architectures for the upgrade circular economy. In the following, five contributions that are particularly representative of the problem under investigation are discussed in detail, as they explicitly consider both product architecture and circular strategies and are therefore especially relevant. A broader comparison of additional contributions is provided in Figure 1.

TCHERTCHIAN ET AL. [29] propose an approach for optimising product architectures to facilitate circular economy design by transforming traditional architectures into modular designs that support industrialised remanufacturing. Products are decomposed into modules that can be remanufactured, recycled, or upgraded, and these modules are grouped using affinity criteria such as reliability, obsolescence, and commonality across product lines. An algorithm classifies modules and evaluates alternatives based on economic and ecological criteria. The resulting modular architectures simplify disassembly and component reuse and provide economic and environmental benefits. However, the approach does not specify how upgrades are systematically integrated into existing product architectures.

POZO ARCOS ET AL. [30] review product architecture design strategies and guidelines that support circular strategies such as repair, refurbishment, and remanufacturing. They derive a categorisation of guidelines for functional product architecture and emphasise considering recovery processes from the earliest design phases. Their recommendations focus on modular designs that allow easy replacement of components, the use of standardised connectors for non-destructive disassembly, and the use and separation of recyclable materials. Products should be designed for upgradability and adaptability to future requirements in order to extend their lifespan. The guidelines strongly support cleaning, diagnosis, disassembly, and reassembly, but they place less emphasis on industrial remanufacturing and upgrade processes in the specific sense of the upgrade circular economy.

YAMADA ET AL. [8] present a decision-support method for designing product architectures with enhanced upgradability. Their approach aims to optimise both the upgrade cycle and the overall product structure and addresses the trade-off between modular and integral architectures. They introduce evaluation models for environmental impact, cost, and customer dissatisfaction, which are combined into an upgradability index to support architectural design decisions. Upgrade considerations are thus embedded early in the design process. The method, however, does not address how upgrades are practically implemented on products or how technological disruptions over time influence the suitability of the chosen architecture.

MATSUDA ET AL. [31] propose a methodology for systematic upgrade planning to extend product lifecycles and reduce waste. The approach uses a component database with information on cost, performance, and market timing to align user demands with product structures and to develop suitable design solutions. Upgrade options are then evaluated across several product generations, linking upgrade planning to market timing and product evolution. Nevertheless, the methodology remains focused on the planning and evaluation of upgrade options and does not show how these upgrades are structurally integrated into the existing product architecture.

SCHUH ET AL. [5] introduce an approach to upgradable product architecture design that is based on modular product structures and circular economy principles. The approach promotes extending product lifecycles through modular upgrades and relies on the standardisation of interfaces and the reduction of interdependencies between modules to create inherently upgradable products. It is closely aligned with circular strategies such as remanufacturing, refurbishing, and recycling and stresses the systematic integration of upgrades into products and their architectures while anticipating technological advancements. However, the approach does not provide a quantitative metric for comparing alternative architectures with respect to their upgrade circularity.

○ not fulfilled ◔ hardly fulfilled ◑ partly fulfilled ◕ largely fulfilled ● fully fulfilled	MATSUDA ET AL. (2003) [31]	TCHERTCHIAN ET AL. (2011) [29]	SCHIFFER (2013) [27]	MESTRE ET AND COOPER (2017) [33]	POZO ARCOS ET AL. (2018) [30]	SHAHBAZI AND JÖNBRINK (2020) [35]	YAMADA ET AL. (2020) [8]	CARLSSON ET AL. (2021) [36]	SCHUH ET AL. (2023) [5]	PSAROMMATIS ET AL. (2024) [34]
Explicit consideration of product architecture design	◑	●	●	◔	●	◑	●	◔	●	◔
Focus on upgradability / upgrade circular economy	●	◕	○	◑	◑	◔	●	◑	●	◑
Market- and technology-driven change requirements	●	◑	●	◔	◔	◔	◑	◑	●	◔
Detailed analysis of modularity and interface flexibility	○	◑	◑	◔	◔	◔	◑	○	◕	○
Quantitative assessment of the architecture's upgrade circularity	◔	◔	◔	○	○	○	●	○	◔	◔

Figure 1: Overview of existing approaches

In summary, numerous scientific approaches address circular economy-oriented product architecture design. The literature shows that modularization and standardized interfaces are key enablers of upgradability and extended product lifecycles, and several methods provide design guidelines, assessment criteria, and optimization frameworks for circular strategies such as reuse, refurbishment, and remanufacturing.

However, significant research gaps remain. Existing approaches remain largely at high-level design principles and do not systematically link product architecture to market- and technology-driven change requirements over the lifecycle. Methods that consider modular architecture rarely offer concrete procedures for assessing interface flexibility and adaptability. This is consistent with the findings of Khan et al. [32], who show in their systematic review of upgradability as a product lifetime extension strategy that, despite its recognized relevance, there is still no widely accepted design framework for upgradable products in a circular economy context. Based on these identified deficits, the framework is derived by operationalizing the core requirements: linking product architecture to external change drivers, enabling a systematic

identification of upgrade-critical elements, and providing a quantitative basis for comparing architectural alternatives. The resulting five-step structure reflects a sequential resolution of these requirements, progressing from architectural foundations to analysis and targeted improvement.

4. Results

The concept for designing upgradeable product architectures in the context of upgrade circular economy is based on five steps. In the first step, product architecture design strategies relevant to the context of upgrade circular economy are identified. In the second step, the interface flexibility of individual product structure elements is determined. Subsequently, the external change variance of the functional structure elements is determined to estimate future upgrades of the product. Based on the results, critical elements of the product architecture can be identified in a fourth step with regard to their upgradeability. Finally, recommendations for action are derived for applying design strategies to the identified critical product architecture elements. While the concept provides a general methodological framework, several analytical procedures, such as the identification of customer features, relevant technologies, or critical elements, must be adapted to the characteristics of the respective product. Accordingly, the later steps of the concept are presented in an application-oriented manner, demonstrating how the methodology can be operationalized and how the results of previous steps feed into targeted architectural improvement. Fig. 2 shows the five steps.

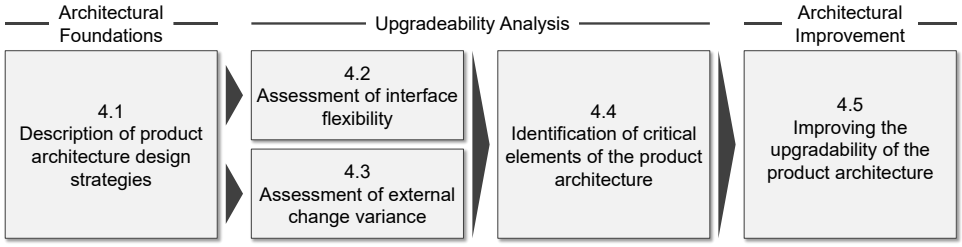


Figure 2: Overview of the concept

4.1 Description of product architecture design strategies in the context of upgrade circular economy

The first step of the concept addresses the systematic identification and structuring of design strategies that enable product architectures to remain upgradeable over multiple life cycles. This step comprises two overarching steps. First, generic product architecture design strategies are identified that influence the functional structure, the product structure, or the transformation relationships between both. To establish this set of strategies, a systematic literature analysis was conducted, from which suitable design principles relevant to upgradeability were derived. In the context of the overarching methodology, only the most essential results of this analysis can be presented here in exemplary form. Future publications will present these findings in more detail. The strategies encompass fundamental principles such as functional decoupling, which reduces dependencies and facilitates future technical changes. In a second step, the identified strategies are examined regarding their conceptual relationships to derive a logical application sequence. This logical sequence becomes relevant when applying the strategies to a concrete product architecture, as certain architectural adjustments structurally precede others. For this purpose, dependencies and mutual influences between strategies are compared and consolidated. The strategies can thus be grouped into broader types of redesign, ranging from structural adjustments that reshape larger architectural segments, to module-level adaptations, and finally to detailed refinements affecting individual interfaces or component properties. These stages represent coherent areas of architectural redesign, each describing a distinct contribution to enhancing upgradeability. By analyzing the relationships between the stages, a defined sequence is derived, which is presented in this paper in exemplary form. The resulting sequence begins with structural and functional decoupling, i.e. defining modular boundaries and reducing interdependencies, followed by over-engineering measures that ensure robustness and enable future

performance. Subsequently, strategies that enhance disassembly and reassembly capability are applied to facilitate upgrade implementation across life cycles. The final stage consists of interface design, enabling standardized and interoperable upgrade connections. The result of this step is therefore a structured set of generic product architecture design strategies, consolidated into overarching stages with a transparent and logically derived order of application.

4.2 Assessment of interface flexibility in product architectures in the context of upgrade circular economy

The second step of the concept addresses the assessment of interface flexibility within product architectures, as flexible interfaces constitute a central requirement for enabling disassembly, re-assembly, and the integration of upgrades throughout the product lifecycle. This step comprises two main steps. First, relevant influencing factors that determine interface flexibility are identified, structured, and operationalized with a specific focus on the de- and re-assembly capability of product architecture elements. In this context, the most relevant influencing factors identified include mechanical, electrical, and software-related separability, accessibility conditions, the degree of interface standardization, and compatibility requirements. In a second step, the operationalized factors are used to derive consistent reference cases that describe distinct types and flexibility levels of interfaces. For this purpose, the characteristics of the influencing factors are consolidated and examined with respect to their combinatorial consistency. Clusters of complementary characteristics form reference cases that represent clearly distinguishable interface variants, each with specific implications for the upgrade potential of the product architecture. Based on this analysis, four representative reference cases for interfaces were derived: (A) a *smart upgrade interface* with high interoperability and automation potential, (B) an a*daptive standard interface* with moderate flexibility due to largely standardized coupling conditions, (C) an *interface with restricted access,* which is characterized by limited physical accessibility and increased installation effort, and (D) an *interface with destructive disassembly*, which requires irreversible separation steps and offers minimal upgrade suitability. Through the structured description of each interface element using the defined influencing factors and their operationalized characteristics, any product interface can now be systematically classified. By comparing the expressed characteristics with those of the reference cases, interfaces can be assigned to one of the four archetypes based on similarity. The result of this step is a catalogue of interface reference cases, evaluated with respect to their flexibility and suitability for future upgrades within the upgrade circular economy.

4.3 Assessment of external change variance of the functional structure elements to estimate future upgrades of the product

The third step of the concept focuses on assessing how strongly individual elements of the functional structure are expected to be affected by future changes arising from evolving customer needs or technological developments. This assessment provides an indication of which functions may require architectural flexibility to accommodate upgrades over multiple usage cycles. The procedure consists of three complementary steps. First, the potential for future changes in customer-relevant features is determined. These customer features are identified individually for the specific product under consideration. To evaluate the likelihood of future changes, a set of systematic assessment criteria was established, forming a framework for determining the change potential of customer features. The criteria include trend dynamics, competitive dynamics, and the Kano classification of customer relevance. By combining change likelihood with feature relevance, a structured assessment of customer feature change potential is derived, distinguishing between *low*, *medium*, and *high* potential. As with customer features, the relevant technology fields are product-specific and are therefore identified individually for the product under consideration. To support this identification, established foresight methods such as technology radars or trend monitoring tools are applied to systematically capture technological developments with potential relevance for the product architecture. Once the relevant fields are identified, each technology is evaluated using a PESTLE-based assessment framework that examines political, economic, social, technological, legal and environmental drivers

influencing the probability of a technological breakthrough. By combining these influencing factors, the overall disruption potential of each technology field is derived and classified into *low*, *medium*, or *high*. In the final step, the results of both analyses are consolidated and mapped onto the functions of the existing product architecture. By linking each function to the customer features and technology fields it depends on the degree to which future market or technology changes may affect the function can be systematically evaluated. The combined assessment allows the functions to be classified according to their expected change variance. Functions with high variance represent areas where architectural flexibility is particularly important in order to support future upgrades and ensure long-term viability within the upgrade circular economy.

4.4 Identification of critical elements of the product architecture with regard to upgradeability

The fourth step of the concept focuses on identifying elements of the product architecture that represent critical constraints for achieving upgradeability across multiple life cycles. As a foundation, the existing product architecture is modelled. For this purpose, the architecture is represented as a graph model that visualizes functions, components, and their interfaces. The first step is the identification of architecture elements whose characteristics indicate limited suitability for future upgrades. For this purpose, the results of the previous steps, the flexibility of interfaces and the change variance of functional elements, are consolidated and mapped onto the established graph architecture model, shown in Fig. 3.

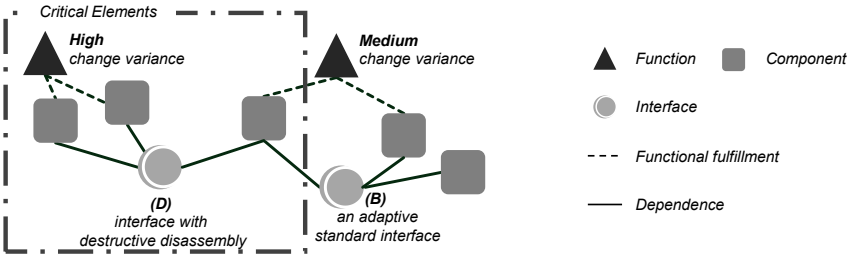


Figure 3: Graph architecture model

By integrating component durability into this analysis, which captures the likelihood that a component requires replacement due to wear-induced degradation, discrepancies become visible between the need for flexibility and the actual flexibility available. Elements are considered critical when they combine high expected change variance or low durability with interfaces that offer only limited upgradeability. Such elements represent potential bottlenecks for future upgrade scenarios and highlight where architectural redesign is most needed. In a second step, the identified discrepancies are used to derive differentiated criticality levels. To this end, the three key assessment dimensions, *functional change variance*, *wear susceptibility of components*, and *interface flexibility,* are combined into a consistent evaluation framework. This enables the classification of architecture elements according to their criticality for future upgrades, see Fig. 4. Highly critical elements are those that face substantial functional or technological change pressure but provide insufficient modularity and interface flexibility to accommodate these changes. The outcome of

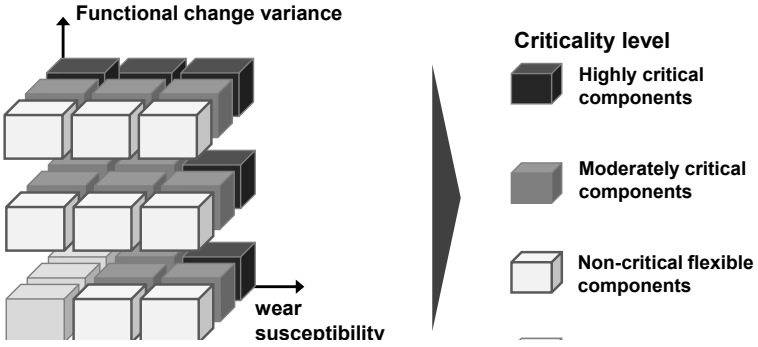


Figure 4: Visualization of the criticality of individual elements

this step is a transparent identification and classification of critical elements within the product architecture, highlighting where targeted design strategies must be applied to ensure long-term upgradeability.

4.5 Improving the upgradability of the product architecture

The final step of the concept focuses on improving the upgradability of the product architecture by systematically applying design strategies to the critical elements identified in the previous step. For each critical architecture element, the most appropriate strategy is then selected based on its individual criticality profile, ensuring that redesign measures respond directly to the specific limitations of the component, function, or interface. In a second step, the selected strategies are applied iteratively to the product architecture. The strategies are implemented in the predefined sequence of stages established in Step 4.1, ensuring that fundamental architectural adjustments are made before more specific refinements. Within each stage, elements are addressed according to their criticality, beginning with highly critical components and progressing to elements of moderate significance.

This approach supports a targeted and efficient improvement of the architecture while maintaining design consistency. The iterative process follows a man-in-the-loop principle, allowing designers to continuously evaluate intermediate results and make context-specific adjustments to the evolving architecture. The third step concerns the evaluation of the improved architecture. To assess the effect of the applied design strategies, the criticality of the updated architecture is quantified using a weighted measure that reflects the distribution of highly critical, critical, and moderately critical elements. The resulting normalized value provides an interpretable indicator of the upgradeability of the revised architecture. By comparing this value to the initial state, it is possible to determine the extent to which architectural bottlenecks have been reduced and the overall suitability for future upgrades has been improved. The outcome of this step is a validated and systematically enhanced product architecture that demonstrates increased upgradeability and supports long-term circular usage through reduced criticality and improved adaptability.

5. Conclusion

This paper presented a concept for designing upgradeable product architectures in the context of the upgrade circular economy. With growing pressure on resources, market dynamics, and technological innovation, companies increasingly require product architectures that remain adaptable over multiple life cycles. The proposed five-step methodology provides a structured approach to integrating upgradeability into early design decisions and systematically evaluating and improving existing architectures. In the first step, generic product architecture design strategies, such as modularization, interface standardization, and functional decoupling, are consolidated and structured into a logical sequence of stages. These stages form the architectural foundation for upgradeability. The second and third steps provide an integrated upgradeability analysis by assessing key determinants of architectural adaptability: the flexibility of interfaces within the product structure and the external change variance of functional elements. Together, these analyses capture both internal architectural constraints and external dynamics arising from evolving customer expectations and technological disruptions. Building on these insights, the fourth step identifies critical elements within the product architecture where low durability, high functional change variance, or insufficient interface flexibility may hinder future upgrades. The final step operationalizes architectural improvement by assigning suitable design strategies to these critical elements and iteratively applying them. The evaluation of architecture alternatives through a quantitative upgradeability evaluation provides a transparent basis for comparing redesign options and guiding decision-making in early development phases. By integrating architectural, technical, market, and technology perspectives into a coherent process, the presented concept enables companies to proactively design product architectures that support modular reassembly, functional extension, and long-term circular value creation. Further refinement of the proposed concept to increase its practicality and transferability to industrial development processes represents an important next step in the ongoing research of the authors. Future research should focus on validating the approach in industrial case

studies, refining the upgradeability evaluation, and exploring the integration of ecological and economic metrics to further strengthen the framework's practical applicability. In addition, the individual methodological steps will be presented and discussed in greater technical depth in future publications to allow a more detailed understanding of their operationalization and implementation.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Institute for European Environmental Policy: European Green Deal Barometer 2024. https://ieep.eu/wp-content/uploads/2025/06/Green-Deal-Barometer-IEEP-2025.pdf. Accessed 19 November 2025.

[2] Alex Colloricchio. The Circular Gap Report 2025. Accessed 19 November 2025.

[3] European Parliament and the Council of the European Union, 2024. Establishing a framework of measures for strengthening Europe's net-zero technology manufacturing ecosystem. Official Journal of the European Union.

[4] Luukkonen, R., Närvänen, E., Becker, L., 2024. Consumer collectives in the circular economy: A systematic review and research agenda. Sustainable Production and Consumption 45, 281–293.

[5] Schuh, G., Kuhn, M., Keuper, A., Patzwald, M., Schenk, L., Feucht, M., Kantelberg, J., Rossmair, G., Schroth, H., Viethen, U., Zeller, P., Guo, D., 2023. Neue Modularität und Technologie-Roadmapping, in: Empower Green Production. Aachener Machine Tool Kolloquium AWK, Aachen, pp. 92–115.

[6] Lloret, A., 2016. Modeling corporate sustainability strategy. Journal of Business Research 69 (2), 418–425.

[7] Shi, Y., Herniman, J., 2023. The role of expectation in innovation evolution: Exploring hype cycles. Technovation 119, 102459.

[8] Yamada, S., Miyajima, S., Yamada, T., Bracke, S., Inoue, M., 2020. Decision Support Method for Upgrade Cycle Planning and Product Architecture Design of an Upgradable Product. IJAT 14 (6), 919–929.

[9] Schulte, A., Maga, D., Thonemann, N., 2021. Combining Life Cycle Assessment and Circularity Assessment to Analyze Environmental Impacts of the Medical Remanufacturing of Electrophysiology Catheters. Sustainability 13 (2), 898.

[10] Blomsma, F., Brennan, G., 2017. The Emergence of Circular Economy: A New Framing Around Prolonging Resource Productivity. J of Industrial Ecology 21 (3), 603–614.

[11] Bocken, N.M.P., Pauw, I. de, Bakker, C., van der Grinten, B., 2016. Product design and business model strategies for a circular economy. Journal of Industrial and Production Engineering 33 (5), 308–320.

[12] MacNeill, A.J., Hopf, H., Khanuja, A., Alizamir, S., Bilec, M., Eckelman, M.J., Hernandez, L., McGain, F., Simonsen, K., Thiel, C., Young, S., Lagasse, R., Sherman, J.D., 2020. Transforming The Medical Device Industry: Road Map To A Circular Economy. Health affairs (Project Hope) 39 (12), 2088–2097.

[13] Unruh, F. von, Mast, J., Irrek, W., 2024. Circular Economy zwischen Ressourcenschonung und Abfallrecycling, in: Büttner, S., Handmann, U., Irrek, W. (Eds.), Transformation zur Circular Economy. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 51–66.

[14] Lammerts, A., 2024. Circular Economy – Idee und Umsetzung, in: Krauss, S., Plugmann, P. (Eds.), Nachhaltigkeit ermöglichen. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 41–55.

[15] Riesener, M., Kuhn, M., Kelbel, N., Schuh, G., 2023. Concept And Exemplary Application Of Industrialized Re-Assembly To An Automotive Use Case In The Context Of Circular Economy. Hannover : publish-Ing.

[16] Hollander, M.C. den, Bakker, C.A., Hultink, E.J., 2017. Product Design in a Circular Economy: Development of a Typology of Key Concepts and Terms. J of Industrial Ecology 21 (3), 517–525.

[17] Schuh, G., Schmitz, S., Gerret, L., Niwar, L., Welsing, M., Calchera, R., 2023. Framework for Circular Production Economy, in: Empower Green Production. Aachener Machine Tool Kolloquium AWK, Aachen, pp. 288–313.

[18] Schuh, G., Mauß, W., Schmitz, S., Adlon, T., Maetschke, J., Neumann, H., Salzwedel, J., Kozielski, S., Luckert, M., Reuter, C., Schmidhuber, M., Witthöft, J., 2023. Green Re-Assembly Upgrade Factory, in: Empower Green Production. Aachener Machine Tool Kolloquium AWK, Aachen, pp. 314–338.

[19] Schuh, G., Schmitt, R.H., Bergs, T., Brecher, C., Eds. (Eds.), 2023. Empower Green Production.

[20] Han, J., Heshmati, A., Rashidghalam, M., 2020. Circular Economy Business Models with a Focus on Servitization. Sustainability 12 (21), 8799.

[21] Schuh, G., Dölle, C., 2021. Sustainable Innovation: Nachhaltig Werte schaffen, 2. Auflage ed. Springer Vieweg, Berlin, Heidelberg, 306 pp.

[22] Rudolf, S., 2013. Produktgerechte Baukastengestaltung. Diss. Apprimus-Verlag, Aachen.

[23] Schuh, G., Riesener, M., 2018. Produktkomplexität managen: Strategien - Methoden - Tools, 3., vollständig überarbeitete Auflage ed. Hanser, München, 273 pp.

[24] Schuh, G., Lenders, M., Nußbaum, C., Rudolf, S., 2012. Produktarchitekturgestaltung, in: Schuh, G. (Ed.), Innovationsmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 115–160.

[25] Göpfert, J., 1998. Modulare Produktentwicklung: Zur gemeinsamen Gestaltung von Technik und Organisation. Deutscher Universitätsverlag, Wiesbaden.

[26] Schuh, G., 2012. Innovationsmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg.

[27] Schiffer, M.G., 2013. Szenariorobuste Produktarchitekturgestaltung. Diss. Apprimus-Verlag, Aachen.

[28] Krause, D., Spallek, J., Blees, C., Kipp, T., 2018. Modulare Produktstrukturierung, in: Rieg, F., Steinhilper, R. (Eds.), Handbuch Konstruktion. Carl Hanser Verlag GmbH & Co. KG, München, pp. 717–741.

[29] Tchertchian, N., Millet, D., Pialot, O., 2011. Modular Grouping Exploration to Design Remanufacturable Products, in: Hesselbach, J., Herrmann, C.S. (Eds.), Glocalized solutions for sustainability in manufacturing : proceedings of the 18th CIRP International Conference on Life Cycle Engineering, Technische Universität Braunschweig, Braunschweig, Germany, May 2nd - 4th, 2011 / Jürgen Hesselbach, Christoph Herrmann, Editors. Springer, Berlin, pp. 413–418.

[30] Pozo Arcos, B., Balkenende, A.R., Bakker, C.A., 2018. Product Design for a Circular Economy Functional Recovery on Focus, in: excellence in design : proceedings of the 15th International Design Conference. DESIGN 2018, Dubrovnik, Croatia. May 21-24. Faculty of Mechanical Engineering and Naval Architecture : Design Society, Zagreb, Glasgow, pp. 2727–2738.

[31] Matsuda, Shimomura, Kondoh, Umeda, 2003 - 2003. Upgrade planning for upgradable product design, in: 2003 EcoDesign 3rd International Symposium on Environmentally Conscious Design and Inverse Manufacturing. 2003. 3rd International Symposium on Environmentally Conscious Design and Inverse Manufacturing - EcoDesign'03, Tokyo, Japan. 12/8/2003 - 12/11/2003. IEEE, pp. 231–234.

[32] Khan, M.A., Mittal, S., West, S., Wuest, T., 2018. Review on upgradability – A product lifetime extension strategy in the context of product service systems. Journal of Cleaner Production 204, 1154–1168.

[33] Mestre, A., Cooper, T., 2017. Circular Product Design. A Multiple Loops Life Cycle Design Approach for the Circular Economy. The Design Journal 20 (sup1), S1620-S1635.

[34] Psarommatis, F., Azamfierei, V., Lindström, J.D., 2024. Circular product design for allowing re-using and re-purposing of products and components: a conceptual framework for circularity. IFAC-PapersOnLine 58 (8), 306–311.

[35] Shahbazi, S., Jönbrink, A.K., 2020. Design Guidelines to Develop Circular Products: Action Research on Nordic Industry. Sustainability 12 (9), 1–16.

[36] Carlsson, S., Mallalieu, A., Almefelt, L., Malmqvist, J., 2021. DESIGN FOR LONGEVITY - A FRAMEWORK TO SUPPORT THE DESIGNING OF A PRODUCT'S OPTIMAL LIFETIME. Proc. Des. Soc. 1, 1003–1012.

Biography

Frederike Hellwig (*1997) studied business administration and mechanical engineering at RWTH Aachen University and has been a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2022, where she has been working as a group lead development management since 2025. Frederike Hellwig focuses on the development of sustainable and circular product architectures and on the management of product complexity in industrial contexts.

Dr.-Ing. Alexander Keuper (*1991) studied mechanical engineering at RWTH Aachen University and has been a research assistant at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen since 2019, where he received his PhD and was appointed head of the Innovation Management Department in 2024. Dr.-Ing. Alexander Keuper works at the intersection of complexity management, development management, and data intelligence to support sustainable innovation under increasing system complexity.

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh (*1958) has the Chair of Production Engineering at RWTH Aachen University and is a member of the Board of Directors of the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen and the Fraunhofer Institute for Production Technology IPT in Aachen. As a scientist and entrepreneur, Prof. Günther Schuh focuses on disruptive innovations, information and production management, and sustainable mobility solutions.

8th Conference on Production Systems and Logistics

From Static To Dynamic Costing: Managing External Volatility and Evaluating Supply Chain Configurations in MTO Manufacturing

Lukas Fleisch, Lukas Budde, Thomas Friedli

Institute for Production and Supply Chain Management/ University of St.Gallen, St.Gallen, Switzerland

Abstract

Manufacturing firms face growing exposure to volatile and uncertain global markets. Particularly in make-to-order (MTO) environments involving complex, large-scale machinery and equipment, long project lead times amplify the impact of external cost fluctuations on competitiveness and profitability. To date, limited attention has been given to how firms can systematically integrate dynamic external market factors into early-stage project costing. This study addresses this gap through an in-depth single case within the Swiss mechanical engineering industry. Using a Design Science Research (DSR) approach, a conceptual framework and cost prediction tool were developed to capture the effects of dynamically changing cost factors - such as wages, exchange rates, and material prices - on distinct cost components, including in-house production, external devices, and service activities. The framework was operationalized in a dynamic cost model employing exponential smoothing (Holt-Winters) time-series forecasting based on monthly historical data. The findings demonstrate that exchange-rate fluctuations and regional cost differentials can materially alter projected project margins and reshape the relative attractiveness of alternative production locations and associated supply chain configurations. This study contributes to the literature on MTO costing and adaptive supply chain configuration under uncertainty by showing how external market volatility can be operationalized within early-stage decision processes. For practitioners, this tool provides early cost transparency and supports the evaluation of alternative supply chain configurations and pricing decisions.

Keywords

Dynamic Cost Forecasting; Project Costing; Volatility; Supply Chain Resilience; Adaptability

1. Introduction

Manufacturing firms increasingly operate under persistent external volatility driven by fluctuating energy prices, exchange rates, material costs, and geopolitical disruptions. Rather than isolated shocks, these dynamics have become structural features of global markets [1], as reflected for example in the Global Supply Chain Pressure Index [2]. This increased volatility poses particular challenges for MTO and long-cycle manufacturing environments where production is triggered only after order confirmation, leading to extended order-to-delivery times compared to ATO or MTS systems [3]. These extended lead times increase exposure to exogenous cost changes between quotation and execution, creating the risk that initial cost assumptions embedded in quotations become outdated and erode margins. In practice, quotation processes frequently rely on fixed parameter assumptions at the time of bidding and do not systematically incorporate evolving input prices, wages, or currency developments. Existing research on supply chain risk and resilience provides valuable guidance for long-term strategic positioning ([4], [5]) but rarely examines how cost-side

DOI: https://doi.org/10.15488/21009
ISSN: 2701-6277

volatility can be operationalized within tactical, project-level manufacturing decisions (e.g., pricing or supply chain selection).

Studies on MTO manufacturing configuration and costing further illustrate this limitation. For example, Li and Womer emphasize the importance of tactical decision-making in MTO networks, using a mathematical optimization model to determine capacity allocation and location decisions [6]. While such approaches offer decision support under predefined cost parameters, they typically assume static or scenario-based inputs rather than dynamically forecasted external cost factors that are embedded in quotation-stage costing processes. Parallel insights from forecasting research – particularly the strong practical performance of exponential smoothing (ETS) models in volatile environments [7] – suggest an opportunity to connect dynamic cost forecasting with project costing. However, the systematic integration of time-series-based external cost forecasts into quotation-stage configuration decisions in long-cycle MTO environments remains underexplored. Therefore, this study addresses the identified gap by examining how dynamically evolving external cost factors can be systematically incorporated into early-stage project costing in long-cycle MTO settings.

The following research question is at the center of this study:

How can dynamically forecasted external cost drivers be systematically integrated into quotation-stage project costing and tactical production configuration decisions in long-cycle MTO environments?

Through an in-depth single case study in the Swiss mechanical engineering industry, following a DSR approach (see section 3), we develop a conceptual framework linking external cost developments (e.g., material prices, wages, exchange rates) to distinct cost components and operationalize the framework through a cost prediction tool using ETS-based forecasting models. The resulting forecasts are embedded into predefined scenarios, enabling the company to assess margin sensitivity and evaluate tactical project-level production location and supply chain configuration decisions under external volatility, a topic of increased importance.

2. Background

2.1 Risks & External Volatility in Global Manufacturing and Supply Chains

Global manufacturing companies increasingly operate under heightened external volatility, uncertainty, and risk. Christopher and Holweg describe the period following the 2008 financial crisis not as a sequence of temporary shocks, but as the beginning of an "era of turbulence," characterized by structurally higher variance in key cost drivers such as energy and raw material prices, exchange rates, and transportation costs [1]. This development is reflected in indices such as the Global Supply Chain Pressure Index, which documents elevated instability across global supply chains [2].

Supply chain management research has developed extensive frameworks on risk and resilience (e.g. [4], [5]), primarily focusing on strategic mitigation and business continuity. However, comparatively little attention has been paid to how recurring cost-side volatility affects operational costing and tactical decision-making.

To systematize different types of risk exposure, Christopher and Peck distinguish risks internal to the firm (e.g., process and control risks), risks external to the firm but internal to the supply chain (e.g., demand and supply risks), and risks external to the entire supply network, often driven by macroeconomic or environmental factors [8]. The present study focuses on this latter category. Prior research has highlighted price uncertainty in input markets and exchange-rate fluctuations as central external risk drivers and discussed mitigation mechanisms such as financial hedging [9]. Broader risk taxonomies, including recent World Economic Forum assessments, further reference among others, tariff changes or macroeconomic inflation as external sources of uncertainty affecting international production networks [10]. While these

contributions clarify the nature and classification of external risks, they offer limited guidance on how recurring external cost volatility can be operationalized within quotation-stage costing or tactical production network configuration decisions.

2.2 Costing in Make-to-Order (MTO) and Long-Cycle Manufacturing Industries

Building on the above discussion, the operationalization gap is especially pronounced in long-cycle MTO industries, where procurement, production, and service activities occur months or even years after the initial bid. In such settings, the customer order decoupling point is positioned upstream, meaning that production is triggered only after order confirmation [3]. As a result, project lead times are extended and cost exposure persists between quotation and execution. During this interval, external cost drivers may change substantially, increasing the risk that static assumptions embedded in project offers become misaligned with realized costs.

In practice, quotation-stage costing often relies on fixed parameter assumptions at the time of bidding. Firms mitigate uncertainty through hedging instruments, contractual clauses, or long-term agreements (e.g., [11], [9]). While these mechanisms reduce exposure, they do not eliminate the temporal sensitivity of project margins to evolving exchange rates, input prices, or inflation, nor do they provide a forward-looking mechanism for systematically anticipating cost developments.

Established costing approaches in manufacturing provide valuable tools for internal cost transparency but only partially address this challenge. Activity-Based Costing (ABC) improves the allocation of overhead and indirect costs by tracing activities to cost objects, enhancing internal cost accuracy [12]. Target Costing emphasizes market-driven cost targets derived from expected sales prices, particularly in product development contexts [13] and Total Cost of Ownership (TCO) broadens the perspective to lifecycle and procurement-related costs, as acquisition, operating, and end-of-life expenses [14]. While these methods enhance cost transparency and strategic alignment, they typically apply assumed or scenario-based cost parameters and do not explicitly incorporate dynamic external market drivers into quotation-stage project calculations.

Li and Womer [6] highlight that tactical supply chain configuration and costing decisions in MTO contexts remain underexplored. Their optimization framework addresses sourcing, scheduling, and configuration decisions but largely assumes stable input prices while external cost dynamics are excluded from the model. Moreover, the study assumes single sourcing and explicitly invites future research to consider multiple location options with distinct cost structures, supplier ecosystems, and lead times.

Taken together, existing research emphasizes resilience, internal cost allocation accuracy, and configuration under static assumptions. The systematic integration of dynamically forecasted external cost drivers into tactical costing and configuration decisions in long-cycle MTO environments remains underexplored.

2.3 Dynamic Cost Models, Forecasting & Network Configuration

Digital technologies increasingly enable advanced supply chain scenario analysis and data-driven decision support ([15], [16], [17]). Furthermore, digital supply chain research emphasizes digital twins, real-time analytics, and advanced scenario and optimization simulation as mechanisms for managing complex production systems and networks [18]. Given the intensified exposure to external volatility (Section 2.1) and the extended cost exposure in long-cycle MTO environments (Section 2.2), it is notable that the integration of dynamic forecasting into project-level costing remains underexplored in the literature.

Generally, a range of analytical approaches addresses uncertainty in cost estimation and supply chain decision-making. Time-series models such as exponential smoothing and autoregressive methods are widely used to forecast economic and price developments ([7]; [19]). In parallel, machine learning (ML) and neural network-based approaches have gained increasing attention in forecasting and intelligent cost analysis. In quotation-stage costing contexts, however, transparency, traceability, and compatibility with established

costing structures represent critical design criteria. Empirical findings suggest that practitioners often favor interpretable approaches over black-box models when cost calculations must remain auditable [20].

Within this context, time-series methods provide a transparent and computationally efficient mechanism for forecasting external cost drivers. ETS models capture level, trend, and seasonality within a state-space framework and have demonstrated robust empirical performance ([19], [7]). Forecasting has also been applied in cost-related environments [21]. However, evidence of structured embedding of forecasted external parameters into quotation-stage costing and tactical configuration decisions remains limited in the literature.

2.4 Identified Research Gap

Across the reviewed literature, key elements of volatility management, costing methodologies, and forecasting techniques are addressed individually. Supply chain risk and resilience research predominantly operates at a strategic level; MTO costing and configuration studies often rely on static or scenario-based cost assumptions; and forecasting research typically focuses on predictive performance rather than structured integration into quotation-stage decision processes.

What remains insufficiently addressed is their intersection: integrating forecasted external cost drivers into project-level costing to inform tactical production location and configuration decisions in long-cycle MTO environments. In particular, the combined consideration of forecast-based margin evaluation and configuration choice has received limited integrated treatment.

Building on this gap, Section 3 outlines our DSR approach to integrate forecasted external cost drivers into quotation-stage project costing and configuration decisions.

3. Research Approach

We applied a Design Science Research (DSR) approach to develop and evaluate a dynamic costing artefact for customer-specific project offers under changing external cost conditions. DSR is particularly suited for solution-oriented Operations Management problems requiring the design and validation of practical artefacts in organizational settings ([22], [23]). It provides a structured framework linking practical relevance, scientific rigor, and iterative artefact development through the relevance, rigor, and design cycles [24]. The relevance cycle anchors the research in a concrete organizational problem. The rigor cycle connects artefact development to existing theoretical knowledge. The design cycle involves iterative building, testing, and refinement of the artefact within its intended environment. Details of these cycles are reflected in the following sections. The empirical inputs guiding the DSR process are summarized in Table 1.

Table 1: Overview of empirical inputs and design activities

Purpose	Data Source	Interaction Format	Duration	Stakeholders involved	Date
Problem statement and understanding	Semi-structured interview	1:1 interview	1h	COO, problem owners from project offering and management accounting	02.2025
Scoping & identification of dynamic cost factors	Workshop with project team of case company	On-site workshop	4h	Project offering and management accounting	03.2025
Evaluation	Completed ranking template (MS Excel) & consolidated prioritization	Structured ranking exercise (asynchronous) followed by alignment meeting	60 min	3 project team members from project offering and management accounting	04.2025
Field testing and solution development	7 historical project cost data sheets (Excel)	Iterative artefact refinement sessions	6x 30-60 min	Project offering	04.2025-08.2025
Evaluation and finalization	Prototype artefact	Presentation & evaluation meeting	1h	Project offering and management accounting	09.2025

Evaluation	Finalized cost prediction tool	Review meeting	1h	COO, problem owners from project offering and management accounting	11.2025
Handover	Final Excel tool & documentation (Word)	Handover meeting	1h	Research team & problem owner from project offering	12.2025

4. Solution Concept - Artefact development

While existing literature addresses elements of volatility, costing, and forecasting individually, the managerial need for integrated dynamic costing solutions in long-cycle MTO contexts remains pronounced. To address this need, we apply the DSR framework of [24] within a single case company in the mechanical engineering industry in Switzerland.

4.1 Relevance Cycle: Problem Framing and Environment Input

The relevance cycle links the research to the application environment: the case company offers large-scale machinery projects with durations of 12–24 months. Representatives from project offering, controlling, and supply chain management identified increasing volatility in exchange rates, material prices, and wage developments as a growing challenge, while costing practices remained largely static.

The organization expressed the need for a dynamic forecasting-based costing solution capable of anticipating cost developments between quotation and execution. Historical and current project quotation data were provided to support artefact development. The first interviews with the experts indicated that there exists a relevant difference of planned and real costs due to dynamic cost effects. As the problem owner from the project offering team reflected: “For years, our business has been very stable and predictable, only recently the dynamics have increased substantially”.

4.2 Rigor Cycle: Knowledge Base and Theoretical Foundations

Building on the literature review (Section 2), three insights guided the artefact design:

1) External volatility has intensified in global supply chains,

2) MTO costing approaches predominantly rely on static assumptions, and

3) Forecasting methods are rarely integrated into quotation-stage costing models.

These insights grounded the development of the dynamic cost calculation tool and motivated the selection of exponential smoothing as a transparent and implementable forecasting method.

4.3 Design Cycle: Artefact Construction, Demonstration, and Evaluation

4.3.1 Artefact Design and Development

Based on the relevance and rigor inputs, we iteratively developed a conceptual framework linking external cost drivers to distinct cost components of MTO projects. Cost drivers were identified through literature and expert workshops and subsequently ranked by company representatives. The prioritized factors in descending order include material prices, exchange rates, tariffs, wage inflation, transportation costs, third-party price adjustments, and delays in payment plans. The framework distinguishes in-house production, external devices, and service activities and links each to relevant cost drivers (Figure 1).

The framework was operationalized in a dynamic cost forecasting prototype implemented in Microsoft Excel. Based on historic data and expert assessment, we have gathered a master list of the relevant components forming the final machine and the usual time in months after project start that the component is produced or purchased and the typical time the installation services occur.

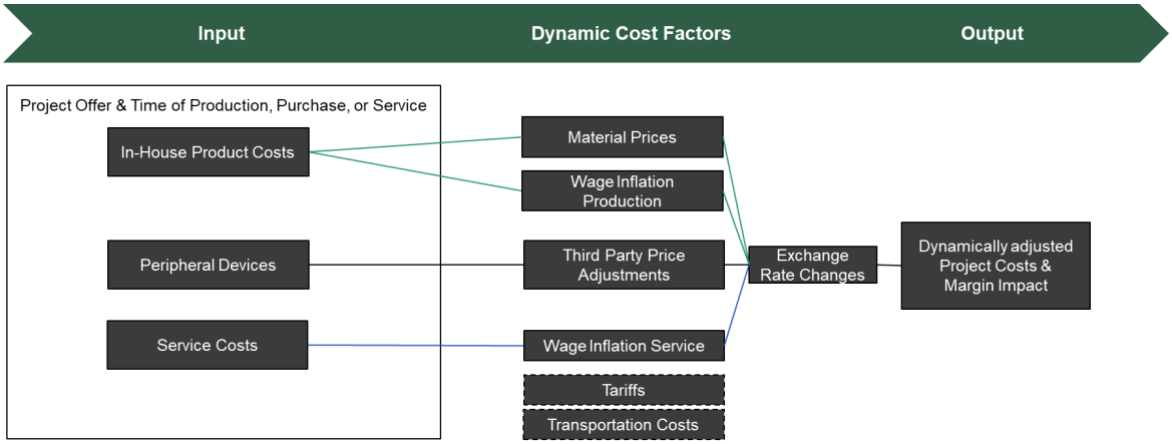


Figure 1: Derived conceptual framework of dynamic project cost calculation

For each selected cost driver (material prices, exchange rates, wage indices), monthly historical time series from January 2022 to September 2025 were collected from publicly available sources. Forecasts were generated using Holt-Winters exponential smoothing (ETS) implemented via MS Excel's FORECAST.ETS function [19], with forecast horizons of up to 24 months, reflecting typical project durations.

Forecast performance was evaluated by comparing out-of-sample forecasts with realized values using Mean Absolute Percentage Error (MAPE). MAPE ranged from 1.16%–4.11% for exchange rates, 2.41%–8.06% for material price indices, and 0.46%–8.41% for wage developments, depending on regional configuration. It is important to note that the objective of the forecasting module is not precise macroeconomic prediction, but systematic parameter updating within quotation-stage cost calculations. Conversely, tariffs and transportation costs were not forecasted but treated as adjustable exogenous scenario parameters, reflecting the company's approach to fix them via contractual agreements.

The artefact integrates forecasted change factors into existing quotation files by adjusting static component costs to their expected values at the time of production, procurement, or service delivery. Forecast parameters are embedded in scenario-based configuration analyses, enabling a transparent decomposition of margin effects into baseline forecast impacts and incremental tariff effects. Figure 2 illustrates the resulting transition from static to dynamic margin.

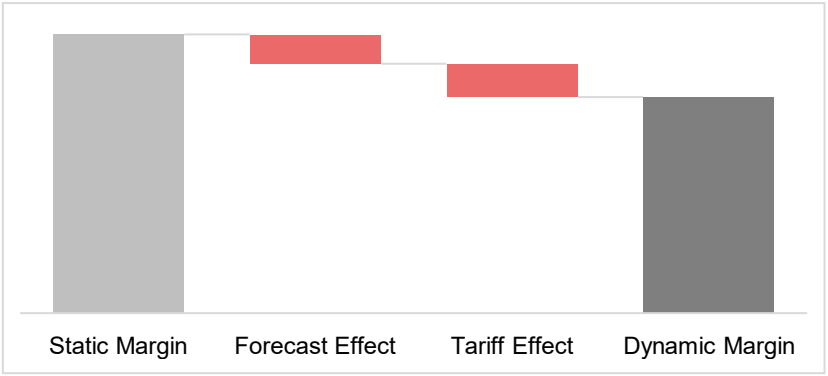


Figure 2: Conceptual transition from static to dynamic margin under forecast and tariff effects

4.3.2 Demonstration

The possible supply chain configurations within the case company include customers in North America, Asia, and Europe, and production in the US, in Switzerland, in the EU, and in China. To illustrate the tactical implications of the developed artefact, a representative project with a 12-month duration from first production to final installation was selected, assuming a project start in October 2025. Motivated by recent

tariff volatility affecting exports to the US, three alternative production configurations were evaluated for a US customer:

In-house production in the US, external devices from US suppliers
In-house production in Switzerland, external devices from European suppliers
In-house production in Europe, external devices from European suppliers

Each configuration was assessed under three tariff regimes:

Baseline: 0% tariff rate on exports from Switzerland and the EU to the US
Scenario 1: 15% tariff rate on exports from both Switzerland and the EU to the US
Scenario 2: 39% tariff rate on Swiss exports and 15% tariff on other European exports to the US

The results are summarized in Table 2.

Table 2: Forecast-based cost and margin simulation results across tariff scenarios

Production Setup	Baseline Cost [Index]	Baseline Scenario: No Tariffs – Forecast Effects Only			15% Tariff EU and CH to US			15% Tariff EU and 39% Tariff CH to US		
		Cost Increase [%]	Margin Loss [%]	Margin Rank	Cost Increase [%]	Margin Loss [%]	Margin Rank	Cost Increase [%]	Margin Loss [%]	Margin Rank
US Production & Peripherals	100.00	1.11	0.79	3	1.11	0.79	1	1.11	0.79	1
CH Production & EU Peripherals	96.68	1.27	0.9	2	12.99	9.21	3	23.57	16.15	3
EU Production & Peripherals	93.49	1.26	0.89	1	12.49	8.86	2	13.36	8.86	2

Under the baseline scenario, the European production configuration yields the highest projected margin, followed by Switzerland, while US production remains structurally more expensive. Forecast-induced cost adjustments alone lead to modest margin deviations below 1%, which nevertheless translate into economically meaningful absolute values for large-scale projects.

Under tariff scenarios, configuration attractiveness shifts substantially. US production becomes relatively more favorable, while Swiss production experiences the strongest margin deterioration under differentiated tariffs. The analysis demonstrates how forecasted cost drivers combined with tariff scenarios systematically inform tactical production location decisions.

4.3.3 Evaluation

The artefact was evaluated through iterative presentations and feedback sessions with project offering, management accounting, supply chain representatives, and the COO. Evaluation criteria included (i) plausibility of forecast adjustments, (ii) relevance for configuration decisions, and (iii) practical applicability within existing quotation processes. Consistent with DSR, evaluation focused on pragmatic validity rather than statistical hypothesis testing ([22], [24]).

Across additional historical projects, forecast-induced margin adjustments ranged between 0.6% and 2.5%. Experts assessed these deviations as realistic and aligned with observed exchange-rate and material price dynamics. The forecasting component was therefore considered sufficiently accurate for structured parameter adjustment within quotation-stage costing.

Beyond numerical plausibility, the artefact enhanced transparency in configuration comparisons and supported more informed evaluation of production timing and supplier-region alternatives under external volatility.

Overall, the evaluation confirmed the artefact's practical utility as a structured decision-support mechanism for managing cost volatility in long-cycle MTO environments.

5. Discussion & Implications

This study contributes to research at the intersection of supply chain volatility, MTO costing, and forecasting. While prior literature addresses external risk and resilience predominantly at a strategic level and MTO configuration and costing largely under static or scenario-based assumptions, the structured integration of time-series-based external cost forecasts into quotation-stage project costing has received limited attention. The contribution of this study does not lie in proposing a novel forecasting method, but in operationalizing established forecasting techniques within a project-level costing and configuration context.

The results demonstrate that forecast-based parameter adjustments, even with moderate accuracy levels (MAPE ranges between approximately 1% and 8%), can materially influence margin assessments when combined with tariff scenarios. By embedding forecasted external cost drivers into quotation-stage calculations and linking them to alternative production configurations, the artefact enables structured tactical decision support in long-cycle MTO environments.

From a managerial perspective, the findings suggest that firms operating in long-cycle MTO settings may benefit from complementing static cost assumptions with transparent, forward-looking parameter updates. The developed artefact illustrates that such integration can be implemented with moderate effort and without requiring complex digital twin architectures or black-box machine learning solutions. Importantly, the forecasting component is not intended as a macroeconomic prediction tool, but as a structured mechanism for updating cost parameters within the typical execution horizon of projects. Its primary value lies in enhancing comparability and transparency of alternative production and sourcing configurations under external volatility.

The study is subject to several limitations. First, the artefact was developed and evaluated within a single case company in the Swiss mechanical engineering industry. Although similar structural characteristics exist in other long-cycle MTO industries, generalizability beyond comparable contexts should be treated with caution. Second, the forecasting component relies on monthly time series and ETS-based methods without incorporating structural breaks or exogenous variables. In highly turbulent environments or under abrupt policy changes, forecast performance may deteriorate. Third, the evaluation follows a pragmatic validation logic consistent with DSR and does not involve statistical hypothesis testing across multiple firms. Future research could extend the approach through multi-case studies, alternative forecasting techniques, or integration into enterprise systems.

6. Conclusion and Outlook

This study examined how dynamically forecasted external cost drivers can be systematically integrated into quotation-stage costing and tactical production configuration decisions in long-cycle MTO environments. By combining established time-series forecasting techniques with scenario-based configuration analysis, the developed framework and cost prediction tool provide a structured mechanism for operationalizing external cost volatility within project-level decision processes.

The findings indicate that even moderate forecast accuracy is sufficient to enhance margin transparency and support tactical configuration choices under exchange-rate and tariff uncertainty. The primary contribution lies in the integration of forecasting and costing within a design-oriented framework.

Future research may further refine the forecasting component, explore hybrid or multi-method approaches, and examine scalability across industries and organizational settings. In increasingly volatile global markets, the structured integration of external cost dynamics into operational decision making remains a promising avenue for both research and practice.

References

[1] Christopher, M., Holweg, M., 2011. “Supply Chain 2.0”: managing supply chains in the era of turbulence. International Journal of Physical Distribution & Logistics Management 41 (1), 63–82.

[2] Benigno, G., Di Giovanni, J., Groen, J., Noble, A., 2022. The GSCPI: A new barometer of global supply chain pressures. Staff Reports 1017. Federal Reserve Bank of New York, New York, NY.

[3] Olhager, J., 2003. Strategic positioning of the order penetration point. International Journal of Production Economics 85 (3), 319–329.

[4] Sheffi, Y. and Rice, J.B., 2005. A supply chain view of the resilient enterprise. MIT Sloan Management Review 47 (1), 41–48.

[5] Pettit, T.J., Fiksel, J., Croxton, K.L., 2010. Ensuring supply chain resilience: Development of a conceptual framework. Journal of Business Logistics 31 (1), 1–21.

[6] Li, H., Womer, K., 2012. Optimizing the supply chain configuration for make-to-order manufacturing. European Journal of Operational Research 221 (1), 118–128.

[7] Makridakis, S., Hyndman, R.J., Petropoulos, F., 2020. Forecasting in social settings: The state of the art. International Journal of Forecasting 36 (1), 15–28.

[8] Christopher, M., Peck, H., 2004. Building the resilient supply chain. International Journal of Logistics Management 15 (2), 1–14.

[9] Bandaly, D., Shanker, L., Kahyaoglu, Y., Satir, A., 2013. Supply chain risk management–II: A review of operational, financial and integrated approaches. Risk Management 15 (1), 1–31.

[10] World Economic Forum, 2025. The Global Risks Report 2025, Geneva.

[11] Tang, C.S., 2006. Perspectives in supply chain risk management. International Journal of Production Economics 103 (2), 451–488.

[12] Cooper, R., Kaplan, R.S., 1988. Measure costs right: make the right decisions. Harvard business review 66 (5), 96–103.

[13] Cooper, R., Slagmulder, R., 1997. Target Costing and Value Engineering. Productivity Press, Portland, OR.

[14] Ellram, L.M., 1995. Total cost of ownership: an analysis approach for purchasing. International Journal of Physical Distribution & Logistics Management 25 (8), 4–23.

[15] Osterrieder, P., Budde, L., Friedli, T., 2020. The smart factory as a key construct of industry 4.0: A systematic literature review. International Journal of Production Economics 221, 107476.

[16] Wang, S., Wan, J., Zhang, D., Li Di, Zhang, C., 2016. Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. Computer Networks 101, 158–168.

[17] Budde, L., Liao, S., Haenggi, R., Friedli, T., 2022. Use of DES to develop a decision support system for lot size decision-making in manufacturing companies. Production & Manufacturing Research 10 (1), 494–518.

[18] Ivanov, D., Dolgui, A., 2021. A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. Production Planning & Control 32 (9), 775–788.

[19] Hyndman, R.J., Koehler, A.B., Snyder, R.D., Grose, S., 2002. A state space framework for automatic forecasting using exponential smoothing methods. International Journal of Forecasting 18 (3), 439–454.

[20] Bodendorf, F., Lutz, M., Michelberger, S., Franke, J., 2022. An empirical investigation into intelligent cost analysis in purchasing. Supply Chain Management: An International Journal 27 (6), 785–808.

[21] Zhao, L., Mbachu, J., Zhang, H., 2019. Forecasting residential building costs in New Zealand using a univariate approach. International Journal of Engineering Business Management 11, 1–10.

[22] van Aken, J.E., 2004. Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules. Journal of Management Studies 41 (2), 219–246.

[23] Holmström, J., Ketokivi, M., Hameri, A.-P., 2009. Bridging practice and theory: A design science approach. Decision Sciences 40 (1), 65–87.

[24] Hevner, A.R., 2007. A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19 (2), 87–92.

Biography

Lukas Fleisch (*1995) is Research Associate & PhD Candidate at the Institute of Production and Supply Chain Management at the University of St.Gallen. He has received a Master's degree in Management, Technology and Economics from ETH Zurich. His main research interests lie in the Management of Global Production Networks.

Lukas Budde (*1985) has been post-doc and division deputy in the field of Production Management at the University of St.Gallen since 2015. His research focuses on smart manufacturing and digital platform ecosystems in manufacturing companies. He is a lecturer for Operations Management courses and for custom programs of the Executive School of the University of St.Gallen. Since 2018, he has also been the managing director of INOS, the largest regional innovation network for SMEs in Switzerland.

Thomas Friedli (*1971) is a professor of production management at the University of St. Gallen. His main research interests are in the field of managing operational excellence, global production management, and management of industrial services. He is a lecturer in the (E)MBA programs in St.Gallen, Fribourg, and Salzburg. Prof. Friedli leads a team of 14 researchers who develop new management solutions for manufacturing companies in today's business landscape. He is also the editor, author, or co-author of 14 books and various articles.

8th Conference on Production Systems and Logistics

Evaluation Of The Realization Effort Of Product Upgrades In The Context Of The Upgrade Circular Economy

Timm Schulz-Isenbeck[1], Alexander Keuper[1], Frederike Hellwig[1], Günther Schuh[1,2]
[1]*Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Aachen, Germany*
[2]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

The manufacturing industry is currently undergoing a fundamental shift away from linear economic models toward a circular economy. A fundamental principle of the upgrade circular economy lies in the extension of product lifecycle, which extends beyond the mere closure of material cycles through recycling. Thus, a variety of measures may be employed to extend the product lifecycle, including reuse, refurbishment, or product upgrades through technical changes. The upgrade circular economy focuses on the prolongation of the product lifecycle through value-enhancing upgrades to preserve resources, increase customer satisfaction and foster sustainability. However, enhancing existing products without modifying their underlaying architectural structure results in reduced flexibility when it comes to incorporating technical modifications and novel features. Thus, the upgradability may be constrained through the existing product architecture, due to the lack of compatibility or outdated interfaces. Companies need to consider both current functional capabilities and potential future adaptability when managing the product lifecycle. This necessitates the evaluation of the realization effort, whether novel features can still be integrated into the existing product architecture. In this regard, it is necessary to consider both the development effort for the upgrade and the re-assembly effort for the implementation into the existing product. This paper presents a holistic overview of criteria to evaluate the realization effort of product upgrades in context of the upgrade circular economy. A methodology is developed to evaluate the realization effort of upgrades for existing products, examining the interdependencies between the scope of the technical change and criteria for development and re-assembly effort.

Keywords

Product Upgrade; Circular Economy; Release Management; Effort Evaluation; Continuous Innovation

1. Introduction

In the current market environment, manufacturing companies encounter significant challenges, primarily due to increasing competitive pressures and a dynamic market environment. Conversely, the finite nature of resources and the observable effects of climate change triggered an increased demand for sustainable products [1]. The circular economy, in contrast to conventional linear economy, is a deliberate and targeted industrial system that prioritizes recovery and renewal [2]. By revisiting conventional linear production models, industries can adopt circular principles to enhance their processes, innovate product offerings, minimize waste, and ultimately secure a competitive advantage in an environmentally aware market [3].

A fundamental principle of the circular economy is the prolongation of product utilization across multiple life cycles. The prolongation can be achieved through a range of strategies, including reuse, refurbishment

DOI: https://doi.org/10.15488/21010
ISSN: 2701-6277

or remanufacturing [4]. Given the prolonged duration of product utilization, it is imperative for manufacturers to guarantee that their products maintain competitiveness and continue to provide value to customers over time [5]. Thus, the implementation of value-enhancing upgrades in the form of technical changes is of particular importance, leading to an upgrade circular economy. The offering of upgrades for existing products enables companies to respond to dynamic market requirements and postpone the obsolescence of the product itself. Additionally, the upgrade circular economy of the products leads to a reduction in resource consumption and an enhancement in the profitability of the products, as upgrades for existing products offer potential for business model innovations. The promotion of modularity leads to the ability to integrate innovations and significantly enhance sustainability [6].

However, the ability to integrate novel features through upgrades might be limited by the product architecture of the existing product generation. Upgrading existing products without modifying their architectural framework results in a reduction of degrees of freedom for the implementation of technical modifications, due to possibly outdated interfaces [7]. This results in the necessity for companies in the upgrade circular economy to evaluate the viability and potential impact of product upgrades as part of the engineering change management.

This paper aims to identify criteria for the evaluation of the realization effort of technical changes in form of product upgrades in context of the upgrade circular economy. Based on the findings, a methodology concept to evaluate the realization effort of upgrades for existing products is developed, examining the interdependence between the scope of technical changes and criteria for development and implementation effort. Thus, the conceptual framework enables systematic assessment of multiple upgrades throughout the product's lifecycle, facilitating release planning within the context of the circular economy of upgrades. First, fundamentals of the upgrade circular economy are presented. Existing approaches from literature that relate to the paper's objective are presented in chapter three. Based on the findings, chapter five is dedicated to the identification of criteria and the subsequent conceptualization of a framework to assess the realization effort.

2. Fundamentals

As the underlying concept, the upgrade circular economy provides the basis for the prolongation of the product utilization through value enhancing upgrades. Technical modifications to existing products are processual supported through the domain of engineering change management, which are presented in the following.

2.1 Upgrade Circular Economy

Manufacturing companies are undergoing a transformation toward a sustainable industry. The concept of sustainability is manifested across three dimensions: economic, ecological, and social [8]. The circular economy serves as the theoretical foundation for this transition from a linear to a circular economic model, thereby underscoring the need for a paradigm shift towards a more sustainable and holistic approach to economic activities [9]. The objective of the circular economy lies in the continuation of raw material cycles. In order to decouple production from finite resources, concrete measures and strategies are required to close the cycles [10].

The overarching strategies of the circular economy are focused on the reduction, closure, and deceleration of raw material cycles. According to POTTING ET AL., the R-strategies assign specific measures to the three strategy dimensions of reduction, closure, and deceleration [11]. In the context of extending product lifetime, the strategy of deceleration is a particularly relevant consideration in this paper. In this dimension, reuse, repair, refurbishment, remanufacturing, and repurpose are possible measures for closing product and material cycles. [4] introduce the term re-assembly to encapsulate the R-strategies. The fundamental principle of re-assembly entails industrial dismantling, value-enhancement, and subsequent re-assembly of products. During this process, value-enhancing technical changes referred to as upgrades can be integrated

by replacing components and modules. Thus, the objective of the re-assembly is not merely to restore the product to its original state, but rather to extend the product's functionality and features. The objective of these upgrades is to maximize customer benefit on a continuous basis. Consequently, this model is characterized as the upgrade circular economy [4].

Therefore, upgradeability is a critical factor in achieving value-enhancing circular products [6]. In this context, the focus for value enhancement lies on capital-intensive products, such as in machinery and plant engineering. To achieve the necessary prolongation of the life cycle and increase the customer value during the usage phase of circular products, it is necessary to ensure that new functionalities and associated technologies can be integrated into the product after it has been produced. Replacing 10-30% of the components of a complex product can increase its value to that of a new product. At the same time, up to 50% of the resources that would have been required for new production can be saved [12]. Against this backdrop, it is particularly important to understand market changes and the respective customers' desires. It is also important to anticipate technological leaps and identify their impact on our products at an early stage [4]. Consequently, when designing an emergent product architecture for the circular economy, future adjustments must be considered.

2.2 Engineering Change Management

The product life cycle can be divided into three main phases. The development cycle includes product planning and development. The production cycle includes production preparation, manufacturing, assembly, and distribution of the product. The usage cycle is the third phase, which includes the use, service, and maintenance of the product, as well as its dismantling and recycling [1].

The characteristics and features of the product are defined during the development cycle [13]. In this phase, it is therefore important to anticipate how the product will be used to design the product accordingly. Thus, about 70% of the lifecycle costs are determined during product development [14]. Concurrently, it is imperative to facilitate adjustments and functional enhancements within the context of the upgrade circular economy. Technical changes can occur during all phases of the product life cycle. However, the scope and impact of technical changes vary depending on the phase. While technical changes in the development cycle are understood as iterations, upgrades represent changes in the usage phase aimed at improving performance [15]. Engineering change management is defined as the systematic process of recording, evaluating, decision-making, documenting, and controlling changes. The objective of engineering change management is to ensure the economic viability of changes and their rapid and reliable implementation [16]. Against this background, the evaluation of the realization effort is a crucial aspect of the viability. The realization effort consists of the development effort of the change and the implementation effort for the re-assembly of the existing product [17]. Accordingly, engineering change management ensures systematic support for the selection and implementation of changes. In this context, it is essential to consider the implications of technical modifications. These can not only affect the product itself, but also have a significant impact on processes, resources, and costs [18]. In view of this, a detailed assessment of the realization effort plays a particularly important role in the planning of product upgrades in the upgrade circular economy.

3. Research Gap

The following section analyses related scientific work regarding the evaluation of the realization effort of product upgrades. Seven relevant approaches in context of this paper are briefly presented and assessed regarding their suitability for evaluating the realization effort of product upgrades within the circular economy.

The work of [19] elaborates design for assembly methods. The authors focus on ease of assembly and found that the application results in simpler, more reliable and less expensive products. The need to respond to the demand for green products and the concomitant need for design for ease of disassembly is highlighted. In

addition to life-cycle design, guidelines for disassembly are described, including some criteria for realization effort evaluation. The case-based study of [20] addresses disassembly and the factors that determine its economic feasibility. Dominant factors influencing the profitability of disassembly operations are identified and used for a sensitivity analysis to determine the economically optimal choice of end-of-life destination. Additionally, the authors present available techniques that determine the productivity of the disassembly. The work of [21] focuses on eco-design that also includes design for disassembly and design for recycling. The authors propose eleven eco-design indicators with corresponding formulas and apply these indicators in a case study of a water source heat pump. The case study examines the connection between the proposed eco-design factors and the traditional factors of a lifecycle impact assessment. [22] elaborates a systematic review to assign green attributes to the product life cycle management phases, namely production, utilization and end-of-life. The results show that attributes of the production lifecycle phase are mostly environmental attributes and often related to production efficiency. The paper concludes that the effects on consumer behavior of most attributes are not examined in studies. In the end, a research agenda is proposed for future developments. In their paper [23] present benefits and challenges for product remanufacturing. The authors conclude through analyses of household appliances and automotive parts that the most important phases of remanufacturing are cleaning and repairing. To ease these phases, they define properties that product designers should focus on during product development. The framework for disassembly metrics of [24] presents a set of requirements for disassembly in the design phase in conjunction with design for environment. Besides some important characteristics of methods and tools, the authors derive ten common factors for disassembly through a literature review. Selective disassembly is defined as a non-destructive operation which is executed by humans. The paper by [25] presents guidelines for product design to ease remanufacturing. Through a literature review, the authors define eight guidelines, categories and properties for product design. In addition, the guidelines are outlined for an integrated remanufacturing system and applied to assess the design of a water meter.

In conclusion, the existing literature highlights approaches to designing environmentally friendly and life cycle-oriented products, with a particular focus on design for assembly, design for disassembly, eco-design, and remanufacturing. These are considered as enabler for the upgrade circular economy, as the necessity arises of the products ability to implement value-enhancing upgrades after the initial production to prolong the product usage. To plan value-enhancing upgrades for existing product architectures it is necessary to determine the respective realization effort. Therefore, the presented approaches emphasize the necessity of assessing the effort of disassembly and re-assembly. However, the existing literature reveals shortcomings in the evaluation of the realization effort of product upgrades. Firstly, none of the approaches provide a holistic overview of criteria to assess the realization effort, considering relevant criteria of both development and implementation effort. Secondly, the approaches focus only on either development or re-assembly, while a holistic assessment of the realization effort for the development and implementation of an upgrade into an existing product is lacking. This results in the underlying research question of this paper, which criteria need to be considered for the evaluation of the realization effort of product upgrades in the context of the upgrade circular economy.

4. Literature Review

To identify the criteria from extant research, a systematic literature review has been conducted following a five step approach based on [26], which is depicted in Figure 1:

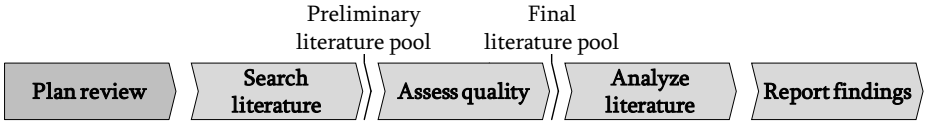


Figure 1: Approach of the literature review based on XIAO AND WATSON [26]

The first step is the planning of the literature review. Therefore, suitable literature databases and search terms have been identified. ScienceDirect was used for literature reviews with various Boolean combinations of keywords and their synonyms as search strings. In the second step the literature search is conducted iteratively across relevant databases and expanded through a snowball search. The identified literature is then checked for access and bibliographic information, cleaned of duplicates, resulting in the preliminary literature pool. The following step compromises the quality assessment of the preliminary literature pool. The preliminary literature pool is then subjected to evaluation regarding its relevance to the research topic. The exclusion of sources is based on several factors, including the date of publication, the use of specific terminology, and the frequency of search terms that are retrieved. The initial screening considers titles and abstracts, while the subsequent, more detailed review also examines the full texts. The outcome of this process is the final literature pool. In the fourth step the literature is systematically examined for criteria to evaluate the realization effort of product upgrades. Thereby a distinction is made between the two dimensions of the development and implementation effort. Each criterion is briefly described in this process. Finally, the identified criteria are consolidated through the merging of similar entries and then categorized through the derivation of mutually exclusive and collectively exhaustive clusters [27]. As a result, the literature review presents a holistic overview of realization effort criteria, forming the basis for a systematic methodology to assess the realization effort of upgrades in context of the upgrade circular economy.

5. Results

The following section presents the results of the literature review. First, general findings are discussed, which are followed by an overview of realization effort criteria in the second subchapter. Finally, a concept for a methodology to assess the realization effort of product upgrades in context of the upgrade circular economy is derived.

5.1 General findings

The preliminary literature pool revealed 3,604 results. Out of the overall results, 32 papers describing criteria for product development, disassembly, re-assembly and remanufacturing criteria were analyzed in detail. After eliminating the duplicates and consolidation of logically redundant criteria, 222 criteria have been identified for further analysis. The criteria were then reviewed for consistency with the scope of this paper and assigned to the two dimensions of the realization effort, namely the development and implementation effort. Thus, 33 criteria for the *development effort* dimension were identified. The *implementation effort* dimension subsumes 25 criteria. The criteria have been clustered to coherent categories with a common evaluation scope. These category clusters are detailed in the following section.

5.1.1 Development effort criteria

The development effort dimension includes all criteria that are related to product development and the planning of product upgrades and changes [1]. The development effort is divided into the categories product complexity, standardization, interfaces, product design and capabilities, which are shown in Table 1 with an excerpt of the respective criteria.

The *product complexity* is primarily determined by the product architecture. This encompasses the evaluation of the number of functions [28], their interdependencies [29] and the complexity of the product structure [19]. Thereby the use of eco-design principles [30] and the avoidance of non-rigid parts [31] needs to be assessed. Also, the number of product features [28] shall be considered to determine the development effort.

Standardization involves the development and implementation of technical standards for product design, ensuring functional requirements and fostering compatibility. Thus, the variance of different components can be reduced. Criteria to evaluate the standardization include the degree of standardization [32], resulting in the multiple reuse of components [19]. Moreover the existence of patents and norms [33], the degree of modularity [25] are criteria to assess the products standardization.

Accompanying the standardization, *interfaces* between components need to be evaluated. To enable the ease the future upgrades, the complexity of interfaces shall be minimized. Therefore, criteria for evaluation are the number of interfaces [19], the type of connection [24] and the quantity of permanent joints [34]. Besides the interface itself, the component compatibility [35] is a crucial criterion in the evaluation of the development effort regarding the interfaces.

Table 1: Categories and excerpt of criteria addressing the development effort

Category	**Criteria**	**Reference**
Product complexity	Number of product functions	[28]
	Product structure	[19]
	Technical feasibility	[36], [33], [37], [38]
	…	…
Standardization	Standardization degree	[39], [25], [29]
	Patents and standards	[33]
	Modular Design	[22], [25]
	…	…
Interfaces	Fastener joint number	[39], [40], [24]
	Standardization of interfaces	[41]
	Component compatibility	[35], [29]
	…	…
Product design	Number of varied materials	[19], [31]
	Material compatibility	[40]
	Design and aesthetics	[22]
	…	…
Capabilities	Product knowledge	[42], [30]
	Employee qualification	[30]
	Access to product data	[39]
	…	…

Ultimately the development effort is influenced by the company's *capabilities*. These reflect overall development and product knowledge [42] or the employee qualification [30]. Moreover the access to product data [39] influences the development effort by drawing conclusions from the usage phase for the development of upgrades. The technical feasibility is closely related to the complexity of the product. For determining the technical feasibility [36] defined capability criteria that help with its estimation, e.g. technology availability [43] and consistency [29].

5.1.2 Implementation effort criteria

Part of the implementation effort represents the second dimension of the realization effort and can be assigned to the production cycle. The dimension includes criteria that characterize the implementation process. The implementation effort is divided into the categories disassembly, assembly and quality control, as presented in Table 2.

The *disassembly* compromises the process of the systematic separation of a product into its components [4]. It is a necessary step to integrate upgrades and novel features into the existing product. Thereby criteria such as disassembly time [44], ease of disassembly [45] or technical difficulty [46] needs to be considered. Also the number of components [30] which need to be removed or modified during the upgrade implementation and their accessibility [40] influence the respective implementation effort. Thus, these factors need to be assessed regarding the implementation effort of an upgrade.

The *re-assembly* is the subsequent step after the disassembly to integrate the upgrade into the product [4]. This step affects the production efficiency and product quality of the upgraded product. Different types of assembly can be differentiated such as manual assembly, automated assembly or hybrid systems varying in the resulting effort for the implementation. The major criterion for addressing the implementation effort regarding the assembly are ease of re-assembly [20]. However, the relations between the subassemblies [35] also influence the implementation effort and thus, shall be considered. These are interlinked with the aforementioned product complexity. Additionally, the direction variability [24] of the assembly process steps need to be considered in the implementation effort evaluation.

Table 2: Categories and excerpt of criteria addressing the implementation effort

Category	Criteria	Reference
Disassembly	Accessibility	[40]
	Ease of disassembly	[45], [39], [40], [20], [41], [22], [30], [47]
	Number of components	[30]
	…	…
Re-Assembly	Ease of Re-Assembly	[20], [41]
	Subassembly relations	[35]
	Direction variability	[24], [23]
	…	…
Quality Control	Ease of detection	[45]
	Ease of cleaning	[19], [20], [41]
	Ease of identification	[23]
	…	…

The *quality control* encompasses all measures to ensure that the product upgrade meets certain quality standards. Thus, it is relevant for both assembly and disassembly, including processes like inspection, testing, evaluation to identify defects and cleaning. Companies can enhance factors like product reliability, customer satisfaction and compliance with industry regulations through quality control. In re-assembly, quality control is necessary in form of final inspections to ensure the products functionality [4]. In regards to the implementation effort evaluation, relevant criteria in terms of quality control are the ease of identification [23] of the component condition, ease of cleaning [41] and ease of detection [45] of components to be replaced.

5.2 Concept for the realization effort evaluation

The results of the presented literature analysis form the basis for a methodology to systematically assess the realization effort of product upgrades in context of the upgrade circular economy. Besides the evaluation of the expected value of upgrades and the future viability of the product architecture, this is an essential step for the release planning in context of the upgrade circular economy [17]. Thus, a concept for the assessment of the realization effort for technical changes in form of product upgrades is outlined below.

The starting point for assessing the realization effort is the scope of the technical change. This can be characterized using a morphological description based on features and characteristics. Relevant characteristics for the change scope are the functional scope, the hardware and software scope and the discipline of the change. The functional scope may encompass the enhancement or preservation of an existing function, as well as the introduction of a new function. In relation to hardware, a distinction can be made between an individual part, an assembly, or a module. Similarly, when describing the scope of the software, a distinction is made between three categories of change: those affecting a single module, those affecting multiple modules, and those affecting the entire system. In addition, it is necessary to provide a detailed description of the scope of the affected hardware and software components. Furthermore, the change

discipline affected by the upgrade is described. The scope of the change may encompass a range of disciplines, including mechanics, electronics, software, or materials. Based on the morphological description of the scope of change, interdependencies to the realization effort can be defined. Therefore, the identified clusters within the development and implementation effort are considered in the schematic procedure, which is visualized in Figure 2.

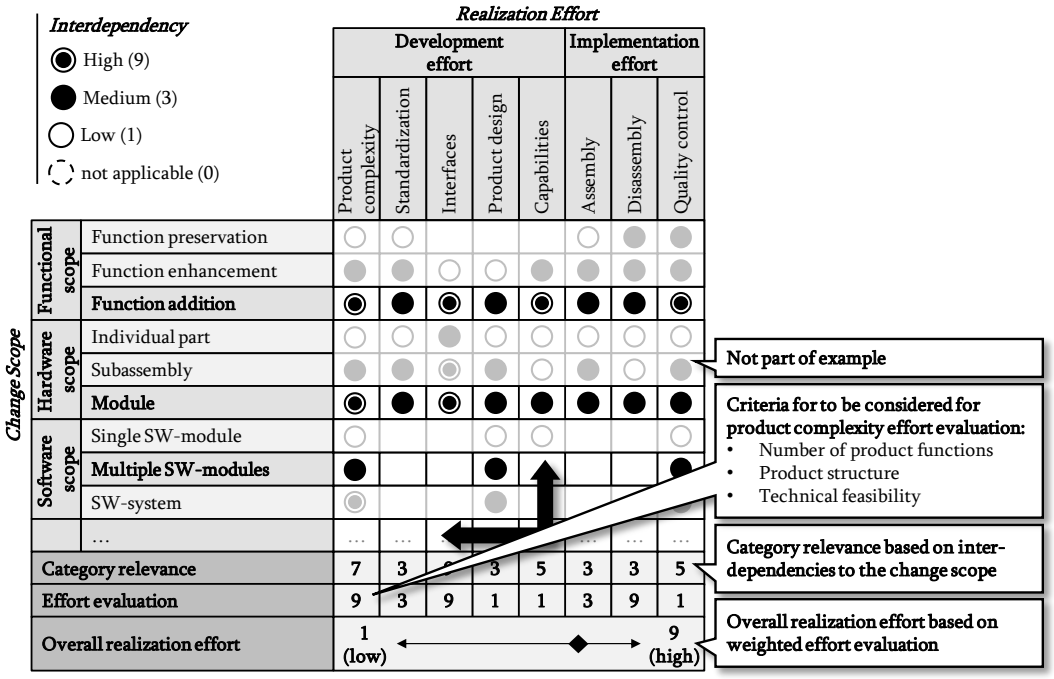


Figure 2: Concept for the evaluation of the realization effort based on td on the change scope

The figure illustrates the assessment of the realization effort required for an upgrade that adds a new function through a replacement of a module and the adaptation of multiple software highlighted in bold fonts. The relevance of the categories considered is based on the analysis of the interdependencies to the change scope, resulting in a weighted distribution of the evaluation criteria. As a result, the realization effort for each development and implementation effort category is determined based on the previously identified criteria in each category. The overall realization effort from the evaluation in the categories weighted according to their relevance. Thus, the conceptual framework enables a systematic assessment of multiple upgrades throughout the product's lifecycle, facilitating the release planning in context of the upgrade circular economy.

6. Conclusion

In summary this paper underscores the necessity to evaluate the realization effort of product upgrades in context of the upgrade circular economy. Thus, not only the development effort of the upgrade itself, but also the implementation effort into the already existing product needs to be considered. Therefore, a systematic literature review has been conducted to identify relevant criteria for the development and implementation effort. The identified criteria have been clustered into overarching categories. Based on these categories a methodology concept has been developed, which enables the evaluation of the realization effort of product upgrades. During this, interdependencies between the change scope and the development and implementation effort are determined. Hence the realization effort can be derived based on the change scope of the upgrade.

Further refinement of the methodology to increase its practicability are the next steps of the research. During this, an extended description of the characteristics of the change scope is planned, to enable a validation of the methodology based on a case study.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 554929445.

References

[1] Schuh G, Dölle C (2021) Sustainable Innovation. Springer Berlin Heidelberg, Berlin, Heidelberg.

[2] Helmold M, Treu J, Fritz J, Hummel F, (Eds.) (2024) ESG, CSR und SDG als langfristiger Wettbewerbsvorteil. Springer Gabler, Wiesbaden, Heidelberg.

[3] Lopes JM, Gomes S, Pacheco R, Monteiro E, Santos C (2022) Drivers of Sustainable Innovation Strategies for Increased Competition among Companies. Sustainability 14(9):5471.

[4] Schuh G, Mauß W, Potente T, Schmitz S, Adlon T, Maetschke J, Neumann H, Salzwedel J, Kozielski S, Luckert M, Reuter C, Schmidhuber M, Witthöft J, :unav (2023) Green Re-Assembly Upgrade Factory. Fraunhofer-Gesellschaft.

[5] Copani G, Behnam S (2020) Remanufacturing with upgrade PSS for new sustainable business models. CIRP Journal of Manufacturing Science and Technology 29:245–56.

[6] Khan MA, Mittal S, West S, Wuest T (2018) Review on upgradability – A product lifetime extension strategy in the context of product service systems. Journal of Cleaner Production 204:1154–68.

[7] Yamada S, Miyajima S, Yamada T, Bracke S, Inoue M (2020) Decision Support Method for Upgrade Cycle Planning and Product Architecture Design of an Upgradable Product. IJAT 14(6):919–29.

[8] Elkington J (1997) Cannibals with forks: The triple bottom line of 21st century business. Capstone, Oxford.

[9] Ranta V, Saari UA (2020) Circular Economy: Enabling the Transition Towards Sustainable Consumption and Production. in Leal Filho W, Azul AM, Brandli L, özuyar PG, Wall T, (Eds.). Responsible Consumption and Production. Springer International Publishing. Cham, pp. 78–89.

[10] McDonough W, Braungart M (2002) Cradle to cradle: Remaking the way we make things. 1st ed. North Point Press, New York, NY.

[11] Potting J, Hekkert MP, Worrell E, Hanemaaijer A (2017) Circular Economy: Measuring Innovation in the Product Chain. PBL Publishers.

[12] Lange U (2017) Ressourceneffizienz durch Remanufacturing - Industrielle Aufarbeitung von Altteilen. Kurzanalyse Nr. 18, Berlin.

[13] Ehrlenspiel K, Meerkamm H (2017) Integrierte Produktentwicklung: Denkabläufe, Methodeneinsatz, Zusammenarbeit. 6th ed. Carl Hanser Verlag, München.

[14] Schuh G, Kuhn M (2016) Lebenszyklusorientierte Produktentwicklung. in Spath D, Westkämper E, (Eds.). Handbuch Unternehmensorganisation. Springer Berlin Heidelberg. Berlin, Heidelberg, pp. 1–17.

[15] Aßmann G (2000) Gestaltung von Änderungsprozessen in der Produktentwicklung. Zugl.: München, Techn. Univ., Diss., 2000. Utz Wiss, München.

[16] Schuh G (2012) Innovationsmanagement. 2nd ed. Springer Berlin Heidelberg, Berlin, Heidelberg.

[17] Riesener M, Keuper A, Schulz-Isenbeck T, Schuh G (2025) Concept for Product Architecture Evolution Planning in the Context of Circular Economy. Procedia CIRP 136:38–43.

[18] Jarratt TAW, Eckert CM, Caldwell NHM, Clarkson PJ (2011) Engineering change: an overview and perspective on the literature. Res Eng Design 22(2):103–24.

[19] Boothroyd G, Alting L (1992) Design for Assembly and Disassembly. CIRP Annals 41(2):625–36.

[20] Duflou JR, Seliger G, Kara S, Umeda Y, Ometto A, Willems B (2008) Efficiency and feasibility of product disassembly: A case-based study. CIRP Annals 57(2):583–600.

[21] Cerdan C, Gazulla C, Raugei M, Martinez E, Fullana-i-Palmer P (2009) Proposal for new quantitative eco-design indicators: a first case study. Journal of Cleaner Production 17(18):1638–43.

[22] Marcon A, Ribeiro JLD, Dangelico RM, Medeiros JF de, Marcon É (2022) Exploring green product attributes and their effect on consumer behaviour: A systematic review. Sustainable Production and Consumption 32:76–91.

[23] Sundin E, Bras B (2005) Making functional sales environmentally and economically beneficial through product remanufacturing. Journal of Cleaner Production 13(9):913–25.

[24] Gkeleri VP, Tourassis VD (2008) A concise framework for disassemblability metrics. IEEE International Conference on Systems, Man and Cybernetics, 2008: SMC 2008 ; 12 - 15 Oct. 2008, Singapore. IEEE. Piscataway, NJ, pp. 1632–1637.

[25] Paul M, Gong SQ-C, Löffler J, Reinhart G (2024) Enhancing product design for remanufacturing: guidelines and their implementation in integrated remanufacturing systems. Production & Manufacturing Research 12(1).

[26] Xiao Y, Watson M (2019) Guidance on Conducting a Systematic Literature Review. Journal of Planning Education and Research 39(1):93–112.

[27] Rasiel EM (1999) The McKinsey way: Using the techniques of the world's top strategic consultants to help you and your business. McGraw-Hill, New York.

[28] Griffin A (1993) Metrics for measuring product development cycle time. J of Product Innov Manag 10(2):112–25.

[29] Sierra-Fontalvo L, Polo-Cardozo J, Maury-Ramírez H, Mesa JA (2024) Diagnosing remanufacture potential at product-component level: A disassemblability and integrity approach. Resources, Conservation and Recycling 205:107529.

[30] Priyono A, Ijomah WL, Bititci US (2015) Strategic operations framework for disassembly in remanufacturing. Jnl Remanufactur 5(1).

[31] Jovane F, Alting L, Armillotta A, Eversheim W, Feldmann K, Seliger G, Roth N (1993) A Key Issue in Product Life Cycle: Disassembly. CIRP Annals 42(2):651–8.

[32] Dziallas M, Blind K (2019) Innovation indicators throughout the innovation process: An extensive literature analysis. Technovation 80-81:3–29.

[33] Dziallas M (2020) How to evaluate innovative ideas and concepts at the front-end? Journal of Business Research 110:502–18.

[34] Saxena V (2025) Enhancing EV battery lifecycle management: Robotic disassembly, design for disassembly, and sustainable solutions. Journal of Energy Storage 119:116368.

[35] Formentini G, Prioli JP, Ko J, Hapuwatte B, Ferrero V, Badurdeen F, Rickli JL, Ramanujan D (2025) A review of disassembly systems for circular product design. Journal of Cleaner Production 506:145459.

[36] Carbonell-Foulquié P, Munuera-Alemán JL, Rodríguez-Escudero AI (2004) Criteria employed for go/no-go decisions when developing successful highly innovative products. Industrial Marketing Management 33(4):307–16.

[37] Hart S, Jan Hultink E, Tzokas N, Commandeur HR (2003) Industrial Companies' Evaluation Criteria in New Product Development Gates. J of Product Innov Manag 20(1):22–36.

[38] Schmidt JB, Sarangee KR, Montoya MM (2009) Exploring New Product Development Project Review Practices *. J of Product Innov Manag 26(5):520–35.

[39] Cappelletti F, Rossi M, Germani M (2022) How de-manufacturing supports circular economy linking design and EoL - a literature review. Journal of Manufacturing Systems 63:118–33.

[40] Aguiar J de, Oliveira L de, Da Silva JO, Bond D, Scalice RK, Becker D (2017) A design tool to diagnose product recyclability during product design phase. Journal of Cleaner Production 141:219–29.

[41] Iberahim F, Abd Wahab D, Mohd Nopiah Z, Abdullah S (2011) Establishment of Remanufacturing Index for Locally Manufactured Automotive Components. KEM 486:77–80.

[42] Caterino M, Iannone R, Macchiaroli R, Riemma S, Pham DT, Fera M (2025) Enhancing remanufacturing operations: A review on Decision-Making models and their implementation challenges. Computers & Industrial Engineering 204:111088.

[43] Carbonell P, Rodriguez Escudero AI, Munuera Aleman JL (2004) Technology Newness and Impact of Go/No-Go Criteria on New Product Success. Marketing Letters 15(2-3):81–97.

[44] Matos J, Martins C, Simões CL, Simoes R (2023) Comparative analysis of micro level indicators for evaluating the progress towards a circular economy. Sustainable Production and Consumption 39:521–33.

[45] Cai Y, Ke C, Ji Q (2024) A Fuzzy Decision-Making Method for Green Design for Remanufacturability. Processes 12(5):911.

[46] Griffin A (2002) Product development cycle time for business-to-business products. Industrial Marketing Management 31(4):291–304.

[47] Yea M, Sheu J-B, Yoo SH (2025) Recycling design and operations in various supply chain structures. Transportation Research Part E: Logistics and Transportation Review 201:104226.

Biography

Timm Schulz-Isenbeck (*1997) studied industrial engineering at RWTH Aachen University. As a research assistant at the WZL Machine Tool Laboratory at RWTH Aachen University, he has been focusing since 2023 on product architecture release planning in the context of the circular economy and the measures for implementation of AI-driven product development.

Dr. Alexander Keuper (*1991) studied mechanical engineering at RWTH Aachen University and has been a research assistant at the WZL Machine Tool Laboratory at RWTH Aachen University since 2019, where he has lead the Innovation Management department since January 2024. In addition to his work at the institute, Dr. Alexander Keuper is the managing director of Complexity Management Academy GmbH.

Frederike Hellwig (*1997) studied industrial engineering at RWTH Aachen University and works as a research assistant at the WZL Machine Tool Laboratory at RWTH Aachen University since 2022, leading the development management team. Her research activities are dedicated to product architecture design in the context of the circular economy.

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh (*1958) is a scientist and entrepreneur who focuses on disruptive innovations, information and production management, and sustainable mobility solutions. Prof. Schuh heads the Chair of Production Engineering at RWTH Aachen University and is a member of the board of directors of the WZL Machine Tool Laboratory at RWTH Aachen University, among other positions.

8th Conference on Production Systems and Logistics

Location of Micro Fulfillment Centers in a Quick Commerce Logistics Network with Consideration of Inventory Management

Zhu Zhu, Haoxun Chen

Optimization Team, Laboratory of Computer Science and Digital Society, University of Technology of Troyes, Troyes, France

Abstract

The rapid growth of e-commerce and the shift in consumer behavior accelerated by COVID-19 have intensified demand for ultra-fast delivery, driving the emergence of quick commerce (Q-commerce). This model relies on a two-echelon structure in which urban Micro Fulfillment Centers (MFCs) receive consolidated shipments from Distribution Centers (DCs) and dispatch orders to nearby customers. While enabling rapid fulfillment, Q-commerce brings challenges in designing and operating a multi-echelon network under uncertain demand, limited storage capacity, and strict delivery-time constraints. This study addresses an integrated Location–Inventory Problem (LIP) for a Q-commerce network composed of DCs, MFCs, and customer zones. The model jointly determines MFC locations, customer-MFC assignment decisions, and product-specific base-stock levels under a periodic-review policy, while allowing reactive lateral transshipments among MFCs to rebalance their inventories in order to reduce stockouts. Demand uncertainty is captured through a scenario-based Mixed Integer Linear Programming (MILP) model with non-anticipativity constraints and a maximum-service-radius restriction reflecting instant-delivery requirements. To efficiently solve the large-scale MILP, we use an enhanced Benders decomposition approach that integrates Pareto-optimal cuts to accelerate its convergence. Numerical experiments validate the proposed model and demonstrate the efficiency of the approach.

Keywords

Supply chain management; Location-inventory; Quick commerce; Micro fulfillment center; Transshipment

1. Introduction

Driven by consumers' increasing preference for immediacy, the rise of instant delivery services has reshaped urban retail logistics. Quick commerce (Q-commerce) represents the most recent evolution in e-tail, it offers delivery of a diverse assortment of daily used products to customers within one hour. Unlike traditional e-commerce that delivers from large warehouses located on the outskirts of a city, Q-commerce relies on a dense network of Micro Fulfillment Centers (MFCs) situated near customers. The proximity of MFCs enables rapid fulfillment but also brings operational complexities to them due to limited storage space, frequent replenishment needs, and significant demand fluctuations.

Despite the importance of MFCs, their location and operation management present several challenges. Existing facility-location models typically assume deterministic demands and do not incorporate inventory dynamics, while classical inventory models rarely consider the spatial configuration of a multi-echelon network. Such separation is insufficient for Q-commerce, where network structure and inventory decisions are inherently intertwined.

DOI: https://doi.org/10.15488/21011
ISSN: 2701-6277

Furthermore, lateral transshipments—an effective method for mitigating stockouts in capacity-constrained inventory systems—remain under-explored in the context of Q-commerce network design. The presence of strict service-distance requirements, coupled with stochastic demand and heterogeneous customer zones, highlights the need for an integrated modeling framework that captures strategic, tactical, and operational decisions in a unified manner.

This study makes two main contributions. Firstly, at the modeling level, we propose a comprehensive framework for jointly considering MFC location and inventory management in a Q-commerce network. Unlike existing studies that treat location and inventory separately, our model simultaneously addresses MFC location, customer-MFC assignment, and inventory control, incorporating lateral transshipments among MFCs to rebalance their inventories under highly variable demand. This integrated modeling approach better captures the multi-echelon and constrained nature of Q-commerce operations.

Secondly, at the solution methodology level, we adopt a scenario-based approach that transforms the considered two-stage stochastic programming problem with stochastic demands into a deterministic equivalent large-scale Mixed Integer Linear Programming (MILP) model. This model considers MFC capacity, lateral transshipments, service-distance constraints, and customer-MFC assignments, covering strategic, tactical, and operational decisions. To efficiently solve the resulting MILP, we use a Benders decomposition algorithm enhanced by Pareto-optimal cuts. This algorithm significantly improves computational efficiency while maintaining solution quality, making it suitable for large-scale instances.

In summary, this study fills a gap in integrated location-inventory modeling for Q-commerce and provides a practical, scalable approach for its network design and operational planning.

2. Literature review

Q-commerce is an emerging model for the retail of foods and daily used products, emphasizing ultra-fast deliveries via micro-fulfillment centers (MFCs or dark stores). Research on Q-commerce is still limited. Most existing studies on Q-commerce are empirical, addressing logistics, urban impacts, and critical success factors [1,2,3,4]. Only few studies have addressed its optimization problems, such as MFC location-routing planning [5], highlighting its network and operational complexities and the need for further research on location and operational planning.

Research on micro-hub/MFC location initially focused on location-only decisions in urban logistics, such as the location of urban LTL (less-than truck) transportation, cross-docking, and postal distribution networks, and hub placement with the consideration of economic, environmental, and social criteria [6,7,8,9,10]. Later studies integrated operational routing decisions with strategical location decisions, in cargo-bike deliveries with zero-emission zones and mixed public-private networks, using metaheuristics and decomposition methods to handle heterogeneous fleets and stochastic demands [11,12,13,14,15,16,17,18]. In Q-commerce, MFC placement has been studied using cluster-first-route-second heuristics and stochastic programming, accounting for multi-period operations and demand uncertainty [17,18]. However, most of these studies focused primarily on facility location and routing, without explicitly considering inventory decisions, limiting their applicability in Q-commerce operations with demand uncertainty.

Research on integrated Location–Inventory Problem (LIP) has evolved from classical facility-location problems by jointly considering supply network design and inventory replenishment decisions. Early studies focused on single-source, two-echelon supply chains (e.g., plant–DC–customer) under commonly used periodic or continuous review inventory policies [19,20,21]. These studies optimize facility opening, customer assignment, order frequency and quantity, incorporating order or inventory-related fixed, replenishment, holding, and transportation costs. Later studies tried to extend LIP to multi-source, multi-product, and multi-period network, or incorporate transshipments between depots or within multi-echelon networks to enhance services and reduce costs [22,23,24]. To capture stochastic demands, approaches such

as chance-constrained programming, Monte Carlo scenario sampling, possibilistic programming, and safety-stock policies have been applied. Puttemans et al. [25] provide a comprehensive review on the integration of strategic (facility location, capacity) and tactical (inventory control, replenishment) decisions. They emphasize the need for more realistic approaches addressing stochastic, multi-period demands and adaptive inventory policies.

From this literature review, we know Q-commerce received limited attention, no paper has addressed location-inventory decisions of MFCs in a Q-commerce network with the consideration of multi-products, maximum service radius restriction, and lateral transshipments among MFCs. This study addresses this gap by jointly considering MFC location, inventory control, distribution center (DC) sourcing, and lateral transshipments in an integrated multi-period and multi-echelon network model with stochastic demands. This model is formulated by adopting a scenario-based approach and solved using an enhanced Benders decomposition approach with Pareto-optimal cuts, enabling efficient optimization of both location and inventory decisions in Q-commerce logistics.

3. Problem description and model formulation

The logistics network considered in this paper has a three-echelon distribution structure consisting of: (i) DCs located on urban outskirts, which consolidate and store products; (ii) MFCs situated within the city, acting as small storage and sorting centers for short-term inventory and rapid order processing and delivery; and (iii) customer zones, which receive products via last-mile delivery from MFCs. In this network, multiple products are transported from DCs to MFCs by trucks, and these products if available and ordered are immediately delivered from MFCs to customer zones using small electric vehicles. This structure reflects practical operations of Q-commerce: decentralized, proximity-based fulfillment that enables ultra-fast delivery. To enhance flexibility and supply security, lateral transshipments between MFCs are adopted to rebalance their inventories in order to reduce their stockout risks.

Building on this network structure, we try to joint optimize MFC location and inventory management in the context of Q-commerce, considering stochastic demands and multiple product types. Key strategic and tactical decisions include MFC location selection, DC-MFC assignments, customer-MFC assignments, and the order-up-to level for each product at each MFC, as well as operational decisions, including lateral transshipment quantities and actual delivery quantities to customers.

To deal with demand uncertainty, we adopt a scenario-based approach to formulate this stochastic optimization problem. A finite set of demand scenarios is generated to represent demand variability across customer zones, products, and time periods, each associated with a probability. Strategic and tactical decisions are made before demand realization in the first stage, while operational decisions—actual deliveries and transshipments—are made in the second stage after demand becomes known. The planning horizon of the second stage is divided into multiple periods, with order delivery in each period. To ensure temporal consistency of the decisions, non-anticipativity constraints are imposed, which prevent the decisions from using the information of future demand realizations not yet observed.

The following assumptions are made to facilitate the model formulation:

- MFCs are created by renovating existing facilities, resulting in heterogeneous storage capacities and operational scales.
- DCs have unlimited storage and processing capacity.
- Each MFC can replenish from only one assigned DC, while transshipments between MFCs can occur from multiple MFCs.
- Each MFC uses a periodic review order-up-to-level policy for the replenishment of each product from a DC, with DC replenishment and MFC transshipment occurring at different times.

- Each MFC has only one inventory replenishment and transshipment opportunity in the multi-period planning horizon of the second stage.
- Customer orders are processed and delivered periodically (in each period) to reduce delivery frequency and improve vehicle capacity utilization.
- Inventory replenishment, transshipment, and order deliveries to customers of MFCs occur at predefined periods, with the inventory level of each MFC reviewed at the end of each period.
- Stockouts incur backorder costs proportional to the demand unfulfilled and the lateness of demand fulfillment.
- Delivery costs linearly depend on the number of vehicles used and their travelling distances.

Based on the problem description and assumptions made above, we introduce the following sets, parameters, and decision variables required for the model formulation.

Sets

N: Set of all nodes in the Q-commerce logistics network considered, with $D \subseteq N$ for DCs, $M \subseteq N$ for candidate MFCs, $C \subseteq N$ for customer zones;

K: Set of products; S: Set of demand scenarios;

T: Set of time periods for order processing and delivery, with $T = \{1,2, \dots, \mathrm{T}\}$, where T is the total number of periods. $T_t \subseteq T$ for transshipments between MFCs;

S_{st} : Set of scenarios that share the same demand history with demand scenario s until period t . $S_{ts} = \left\{l \in S \backslash \{s\} \colon d^l_{ik\tau} = d^s_{ik\tau}, \forall i \in C, \forall k \in K, \forall \tau \in [1,t]\right\}$, where $d^s_{ik\tau}$ is the demand of customer zone i for product k in period τ under scenario s.

Parameters

p_s: Probability of demand scenario s, $\forall s \in S$;

l_{ij}: Distance from node i to j, $\forall i, j \in N$; l_{max}: Maximum service radius (distance) of each MFC;

v_k: The volume per unit of product k, $k \in K$; V^i_{max}: Maximum storage volume of MFC i, $\forall i \in M$;

c_i: Creation cost of MFC i, $\forall i \in M$.

c_r, c_t, c_v: Unit transportation cost per unit of product per km from DCs to MFCs, between MFCs, and from MFCs to customers, respectively;

c_k: Unit handling cost per unit of product k, $k \in K$.

h_{ik}, b_k: Holding cost and backorder cost per period per unit of product k for MFC i, $\forall i \in M$, $\forall k \in K$.

μ_{ik}: Expected demand of product k for customer zone i in the planning horizon, $\forall i \in C, \forall k \in K$.

d^s_{ikt}: Realized demand of product k for customer zone i in period t under scenario s, $\forall i \in C, \forall k \in K, \forall t \in T, \forall s \in S$.

Decision Variables

x_i: Binary variable, $x_i = 1$ if MFC i is opened, otherwise $x_i = 0$, $\forall\, i \in M$.

y_{ij}: Binary variable, $y_{ij} = 1$ if MFC i replenishes from DC j, otherwise $y_{ij} = 0$, $\forall\, i \in M, \forall j \in D$.

z_{ij}: Binary variable, $z_{ij} = 1$ if MFC i serves customer zone j, otherwise $z_{ij} = 0$, $\forall\, i \in M, \forall j \in C$.

o_{ik}: Real variable indicating the order-up-to level of product k at MFC i, $\forall\, i \in M, \forall k \in K$.

q^s_{ijkt} : Real variables indicating the quantity of product k delivered from MFC i to customer zone j in period t under scenario s, $\forall\, i \in M, \forall j \in C, \forall\, k \in K, \forall t \in T, \forall s \in S$.

g^s_{ijkt}: Real variables indicating the quantity of product k transshipped from MFC j to MFC i in period t under scenario s, $i \neq j$, $\forall\, i \in M, \forall j \in M, \forall\, k \in K, \forall t \in T, \forall s \in S$.

Auxiliary Variables

r^s_{ijk}: Quantity of product k replenished from DC j to MFC i under scenario s, $\forall\, i \in M, \forall j \in D, \forall\, k \in K, \forall s \in S$.

I^s_{ikt}: Inventory level of product k at MFC i at the end of period t under scenario s, $\forall\, i \in M, \forall\, k \in K, \forall t \in T, \forall s \in S$.

I^s_{ikT}: Inventory level of product k at MFC i at the end of the final period under scenario s, $\forall\, i \in M, \forall\, k \in K, \forall s \in S$.

OH^s_{ikt}, BO^s_{ikt}: On-hand inventory level and backorder level of product k at MFC i at the end of period t under scenario s, $\forall\, i \in M, \forall\, k \in K, \forall t \in T, \forall s \in S$.

By considering a finite number of demand scenarios defined by S, the integrated MFC location and inventory problem considered is formulated as a MILP model with the objective of minimizing a total cost. This total

cost contains four main components: (i) MFC configuration costs, including creation and operational expenses; (ii) transportation costs, involving DC-to-MFC transportations, MFC-to-customer deliveries, and inter-MFC transshipments; (iii) inventory costs, including both holding and backorder costs at MFCs; and (iv) customer order processing costs at MFCs. The full formulation of the model is given below.

$$\begin{aligned}\min & \sum_{i\in M} c_i \cdot x_i + \sum_{s\in S} p_s \Big(\sum_{i\in M}\sum_{j\in D}\sum_{k\in K} c_r \cdot l_{ij} \cdot r_{ijk}^s + \sum_{i\in M}\sum_{j\in C}\sum_{k\in K}\sum_{t\in T} c_v \cdot l_{ij} \cdot q_{ijkt}^s \\ & + \sum_{i\in M}\sum_{j\in M}\sum_{k\in K}\sum_{t\in T_t} c_t \cdot l_{ij} \cdot g_{ijkt}^s + \sum_{i\in M}\sum_{k\in K}\sum_{t\in T} h_{ik} \cdot OH_{ikt}^s + b_k \cdot BO_{ikt}^s \\ & + \sum_{i\in M}\sum_{j\in C}\sum_{k\in K}\sum_{t\in T} c_k \cdot q_{ijkt}^s \Big)\end{aligned} \tag{1}$$

s.t.

$$\sum_{k\in K} v_k \cdot o_{ik} \le V_{\max}^i \cdot x_i, \forall i \in M \tag{2}$$

$$\sum_{k\in K}\sum_{j\in C} \mu_{jk} \cdot v_k \cdot z_{ij} \le V_{\max}^i \cdot x_i, \forall i \in M \tag{3}$$

$$I_{ikt}^s \le o_{ik}, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{4}$$

$$\sum_{j\in D} y_{ij} = x_i, \forall i \in M \tag{5}$$

$$\sum_{i\in M} z_{ij} = 1, \forall j \in C \tag{6}$$

$$z_{ij} \le x_i, \forall i \in M, \forall j \in C \tag{7}$$

$$\sum_{j\in M} g_{ijkt}^s \le (V_{max}^i / v_k) \cdot x_i, \forall i \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{8}$$

$$\sum_{i\in M} g_{ijkt}^s \le (V_{max}^j / v_k) \cdot x_j, \forall j \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{9}$$

$$o_{ik} - I_{ikT}^s - V_{max}^i \cdot (1 - y_{ij}) \le r_{ijk}^s \le o_{ik} - I_{ikT}^s + V_{max}^i \cdot (1 - y_{ij}), \forall i \in M, \forall j \in D, \forall k \in K, \forall s \in S \tag{10}$$

$$l_{ij} \cdot z_{ij} \le l_{max}, \forall i \in M, \forall j \in C \tag{11}$$

$$I_{ik1}^s = o_{ik} - \sum_{j\in C} d_{jk1}^s \cdot z_{ij}, \forall j \in M, \forall k \in K, \forall s \in S \tag{12}$$

$$I_{ikt}^s = I_{ik(t-1)}^s + \sum_{j\in M} g_{ijkt}^s - \sum_{j\in M} g_{jikt}^s - \sum_{j\in C} d_{jkt}^s \cdot z_{ij}, \forall i \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{13}$$

$$I_{ikt}^s = I_{ik(t-1)}^s - \sum_{j\in C} d_{jkt}^s \cdot z_{ij}, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, t \notin T_t, \forall s \in S \tag{14}$$

$$OH_{ik1}^s = o_{ik} - \sum_{j\in C} q_{ijk1}^s, \forall i \in M, \forall k \in K, \forall s \in S \tag{15}$$

$$OH_{ikt}^s = OH_{ik(t-1)}^s + \sum_{j\in M} g_{ijkt}^s - \sum_{j\in M} g_{jikt}^s - \sum_{j\in C} q_{ijkt}^s, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{16}$$

$$OH_{ikt}^s = OH_{ik(t-1)}^s - \sum_{j\in C} q_{ijkt}^s, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, t \notin T_t, \forall s \in S \tag{17}$$

$$BO_{ik1}^s = \sum_{j\in C} d_{jk0}^s \cdot z_{ij} - \sum_{j\in C} q_{ijk1}^s, \forall i \in M, \forall k \in K, \forall s \in S \tag{18}$$

$$BO_{ikt}^s = BO_{ik(t-1)}^s + \sum_{j\in C} d_{jkt}^s \cdot z_{ij} - \sum_{j\in C} q_{ijkt}^s, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, \forall s \in S \tag{19}$$

$$\sum_{\tau\in T, \tau\le t} q_{ijkt}^s \le \sum_{\tau\in T, \tau\le t} d_{jk\tau}^s \cdot z_{ij}, \forall i \in M, \forall j \in C, \forall k \in K, \forall t \in T, \forall s \in S \tag{20}$$

$$\sum_{j\in C} q_{ijkt}^s \le OH_{ik(t-1)}^s, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, \forall s \in S \tag{21}$$

$$OH_{ikt}^s \ge I_{ikt}^s, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{22}$$

$$BO_{ikt}^s \ge -I_{ikt}^s, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{23}$$

$$q_{ijkt}^s = q_{ijkt}^l, \forall l \in S_{ts}, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{24}$$

$$g_{ijkt}^s = g_{ijkt}^l, \forall l \in S_{ts}, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{25}$$

$$x_i, y_{ij}, z_{ik} \in \{0,1\}, \forall i \in M, \forall j \in D, \forall k \in C \tag{26}$$

$$o_{ik}, r_{ijk}^s, OH_{ikt}^s, BO_{ikt}^s \ge 0, \forall i \in M, \forall j \in D, \forall k \in K, \forall t \in T, \forall s \in S \tag{27}$$

$$q_{ijkt}^s \ge 0, \forall i \in M, \forall j \in C, \forall k \in K, \forall t \in T, \forall s \in S \tag{28}$$

$$g_{ijkt}^s \ge 0, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{29}$$

$$I_{ikt}^s \in R, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{30}$$

Constraints (2) and (3) ensure that the capacity limit of each MFC is not violated. Constraints (4) restrict the inventory level of each product at an MFC to below or at its corresponding order-up-to level. Constraints (5) imply that each MFC can only replenish from at most one DC. Constraints (6) and (7) ensure that each customer zone is assigned to exactly one opened MFC. Constraints (8) and (9) specify that transshipments are allowed only between opened MFCs and must not exceed their capacity limits. Constraints (10) define the replenishment quantity r_{ijk}^{s} for any i, j, k, s. Constraints (11) impose the maximum service distance requirement for customer-MFC assignments. Constraints (12)-(14) describe the inventory balance of each product at each MFC. Constraints (15)-(17) define all on-hand inventory levels. Constraints (18) and (19) specify all backorder quantities. Constraints (20) and (21) ensure that the quantity delivered by each MFC to its customer zones does not exceed its available inventory and its customer demand to fulfill for each product. Constraints (22) and (23) further bound on-hand and backorder levels. Constraints (24) and (25) are non-anticipativity constraints that guarantee temporal consistency of the decisions of the model. Constraints (26) to (30) specify the domains of all variables.

4. Solution Approach

The MILP model (1)-(30) formulated in the last section is a deterministic equivalent model of the integrated MFC location and inventory problem we consider. It includes strategic MFC location decisions tactical DC-MFC assignment, customer-MFC assignment and MFCs' order-up-to level decisions, and operational inventory replenishments, lateral transshipments, and order delivery decisions of MFCs. The size of the model grows rapidly with the number of candidate MFCs, the number of customer zones, the number of products, the number of time periods, and the number of demand scenarios, while the inclusion of non-anticipativity constraints further increases the computational complexity of the model. Consequently, direct solution of the model using a commercial solver becomes very time consuming for an instance of practical size. To efficiently tackle this model, we adopt a Benders decomposition approach enhanced by Pareto-optimal cuts.

This approach decomposes the original MILP model into a master problem which makes strategic and tactical decisions (MFC location, DC–MFC assignment and customer–MFC assignment, and MFCs' order-up-to levels) in the first stage, and a subproblem which makes multi-period operational decisions (inventory replenishments, lateral transshipments, and order delivery decisions of MFCs) in the second stage. The decisions of the subproblem in each period are conditional on realized demands until this period. More specifically, decision variables x_i, y_{ij}, z_{ij}, o_{ik} are included in the master problem, while decision variables q_{ijkt}^{s}, g_{ijkt}^{s} are assigned to the subproblem. The Benders master problem and subproblem are presented in Appendix 1, where the dual of the subproblem is given in Appendix 2.

Benders decomposition (BD) is an iterative approach, which generates a cut (valid constraint) for the master problem based on the solution of the dual subproblem. In each iteration, the resolution of the master problem determines the values of the decision variables for the first stage, which are then fixed in the subproblem. The resolution of the dual subproblem generates an optimality cut or a feasibility cut for the master problem depending on whether the dual subproblem has an optimal solution or is unbounded. These cuts are incorporated into the master problem. In each iterative, the optimal objective value of the master problem provides a lower bound and the optimal objective value of the subproblem provides an upper bound for the objective value of the original MILP. This alternative resolution of the master problem and the dual subproblem iterates until the relative gap between the upper bound and lower bound falls below a prespecified tolerance, ensuring convergence to an optimal solution or a solution very close to optimum.

To accelerate the convergence of BD, we adopt Pareto-optimal Benders cuts generated by solving an auxiliary subproblem (Magnanti & Wong [26]). These non-dominated cuts provide a tighter lower bound for the master problem at each iteration, thereby reducing the number of iterations required to reach

optimality. Specifically, a Pareto-optimal cut maximizes the objective value of the dual subproblem when the variables of the master problem are fixed to an interior point (called core point) of its convex hull. This technique is particularly effective for instances with tightly coupled location decisions and inventory decisions, where BD with standard Benders cuts converges slowly due to strong interdependencies between strategic and operational decisions.

The procedure of the BD algorithm, including all steps and cut generations, is outlined in Table 1.

Table 1: Procedure of the Enhanced Benders Decomposition Algorithm

Step	Description
0	Construct an initial solution of the MILP and compute the core point for Pareto-optimal cuts
1	Solve the master problem to determine the values of strategic and tactical decision variables (x_i, y_{ij}, z_{ij}, o_{ik}).
2	Solve the subproblem with fixed values of (x_i, y_{ij}, z_{ij}, o_{ik}) to determine the values of operational decision variables (q^s_{ijkt}, g^s_{ijkt}).
3	If the dual subproblem has an optimal solution, generate a Pareto-optimal cut by solving an auxiliary subproblem; otherwise, generate a feasibility cut if the dual subproblem is unbounded.
4	Add the generated cut to the master problem and update the best upper bound and lower bound.
5	Iterate Steps 1–4 until the relative gap between upper bound and lower bound falls below a prespecified error tolerance.

5. Numerical experiments

The enhanced Benders decomposition algorithm with Pareto-optimal cuts is implemented in C++ and solved using IBM ILOG CPLEX 22.1. Experiments are performed on a workstation with an Intel Core i7-11850H (8 cores, 2.50 GHz) and 32 GB RAM, running Windows 10 (64-bit).

Computational experiments are conducted on three sets of instances representing urban Q-commerce networks of varying sizes: (i) 2 DCs, 5 candidate MFCs, 20 customer zones, 2 products, 2 or 4 periods, 100/200/500 demand scenarios; (ii) 3 DCs, 10 candidate MFCs, 50 customer zones, 3 products, 2 or 4 periods, 100/200/300 scenarios; (iii) 4 DCs, 20 candidate MFCs, 100 customer zones, 3 products, 2 or 4 periods, 100/200/300 scenarios. Each configuration is tested with five instances.

The DCs, candidate MFCs, and customer zones are spatially dispersed to reflect realistic urban layouts. The demand scenarios are generated using a scenario tree, with each scenario generated via Monte Carlo sampling to capture stochastic variability. This approach captures spatial and temporal variability of customer demands in Q-commerce. Model parameters, including MFC creation costs, handling, replenishment, and transportation costs, MFC capacities, and inventory-related costs, are calibrated based on recent literature, industry reports, and publicly available data from major Q-commerce providers. Holding and shortage costs are aligned within empirically validated ranges.

The model and algorithm were first validated on small instances (2 DCs, 5 MFCs, 20 customer zones) with known optimal solutions, prior to experiments on larger and more realistic instances. Computational results demonstrate the efficiency and scalability of the enhanced Benders decomposition. The error tolerance in the BD algorithm was set to 1e-4(0.01%) by default, for this setting, most instances could be solved to optimality in a CPU time no more than 600 seconds. However, for the set of largest instances, when the number of scenarios reaches 200, the computational time of the algorithm increases significantly. For this reason, the error tolerance was set to 1e-2(1%) for this set of instances. The small instances with up to 500 scenarios were solved within 80 seconds, while the large instances (up to 4 DCs, 20 MFCs, 100 customer zones) required a maximum CPU time of 4000 seconds. Table 2 reports solution times and optimality gaps across the instances of different sizes, where $|T|$ and $|S|$ represent the number of periods and the number of scenarios respectively. Table 3 examines the effect of the non-anticipativity constraints on computational performance of the BD algorithm, where the Objective Gap is calculated as $\frac{Obj\ without\ NAC - Obj\ with\ NAC}{obj\ with\ NAC} \times 100\%$. The results indicate that the inclusion of non-anticipativity constraints has no significant impact on the objective value.

Table 2: Computational performance of the enhanced Benders decomposition algorithm on different instance sets

Instance Set (DC-MFC-C-K)	$\|T\|$	$\|S\|$	Average CPU Time	Maximum CPU Time	Average Relative Gap	Maximum Relative Gap
2-5-20-2	2	100	1.54 s	1.93 s	0.00%	0.01%
		200	3.49 s	4.33 s	0.00%	0.00%
		500	12.81 s	19.40 s	0.00%	0.00%
	4	100	6.77 s	7.64 s	0.00%	0.00%
		200	17.43s	19.49 s	0.00%	0.00%
		500	77.95 s	96.31 s	0.00%	0.01%
3-10-50-3	2	100	15.65 s	18.50 s	0.00%	0.00%
		200	34.85 s	40.70 s	0.00%	0.00%
		300	63.22 s	73.73 s	0.00%	0.00%
	4	100	85.02 s	97.10 s	0.00%	0.00%
		200	214.81 s	260.53 s	0.00%	0.01%
		300	440.24 s	561.95 s	0.00%	0.01%
4-20-100-3	2	100	104.34 s	122.52 s	0.00%	0.01%
		200	270.67 s	332.56 s	0.01%	0.01%
		300	505.01 s	618.16 s	0.00%	0.01%
	4	100	693.44 s	849.77 s	0.01%	0.01%
		200	1718.6 s	2042.55 s	0.54%	0.71%

Table 3: Impact of the Non-Anticipativity Constraints (NAC) on Computational Performance

Instance Set (DC-MFC-C-K-T-S)	Average Objective Gap	Maximum Objective Gap
2-5-20-2-2-200	0.00%	0.00%
3-10-50-3-2-200	0.00%	0.00%
4-20-100-3-2-200	0.00%	0.00%

6. Conclusion

This study addresses a Q-commerce MFC location and inventory problem with stochastic demands, multiple products, maximum-service radius restriction, and lateral transshipments between MFCs. We propose an integrated multi-period, multi-echelon model that simultaneously considers MFC location, DC sourcing, order-up-to level inventory polices, and inter-MFC transshipments. This stochastic optimization problem is formulated as a deterministic equivalent MILP by adopting a scenario-based approach. To solve this large-scale model, we employ an enhanced Benders decomposition algorithm with Pareto-optimal cuts, where strategic and tactical decisions are optimized in the master problem, and operational decisions in multiple periods are optimized in the subproblem.

The correctness of the model has been validated on small instances, and the algorithm can efficiently solve large instances with networks up to 4 DCs, 20 MFCs, and 100 customer zones. Preliminary analysis indicates that the non-anticipativity constraints have slight effect on solution quality, suggesting that scenario-wise subproblem decomposition may be adopted to further improve the scalability of the algorithm. Potential extensions may consider alternative inventory policies, direct DC-to-customer deliveries, alternative MFC fulfillment strategies, dynamic routing for MFC-to-customer deliveries, and demand forecasting to better meet the immediacy requirements of Q-commerce.

Acknowledgement

The research in this paper was partially funded by the China Scholarship Council (CSC) with the scholarship No. 202406280078 for the PhD study of the first author at the University of Technology of Troyes.

Appendix 1 Benders Master Problem and Subproblem

Let x_i, y_{ij}, and z_{ij} denote the first-stage binary decision variables, representing the opening of MFCs, the replenishment links between DCs and MFCs, and the assignment of customer zones to MFCs, respectively. The continuous variable θ is introduced to represent the optimal value of the subproblem. Let X denote the vector of all first-stage binary variables. Let B denote the coefficient matrix of the original subproblem, and b denote the right-hand-side (RHS) vector of its constraints.

For the dual subproblem, u_i denotes an extreme point when it is bounded, where $i \in I$, I is the set of all extreme points. r_j denotes an extreme ray when the dual subproblem is unbounded, where $j \in J$, J is the set of all extreme rays.

The master problem is formulated as follows:

$$\min \sum_{i\in M} c_i \cdot x_i + \theta \tag{31}$$

s.t.

$$\theta \geq (b - BX)^T u^i, \forall i \in I \tag{32}$$

$$(b - BX)^T r^j \leq 0, \forall j \in J \tag{33}$$

$$\sum_{k\in K} v_k \cdot o_{ik} \leq x_i \cdot V^i_{max}, \forall i \in M \tag{34}$$

$$\sum_{k\in K}\sum_{j\in C} \mu_{jk} \cdot v_k \cdot z_{ij} \leq x_i \cdot V^i_{max}, \forall i \in M \tag{35}$$

$$\sum_{j\in D} y_{ij} = x_i, \forall i \in M \tag{36}$$

$$\sum_{i\in M} z_{ij} = 1, \forall j \in C \tag{37}$$

$$z_{ij} \leq x_i, \forall i \in M, \forall j \in C \tag{38}$$

$$l_{ij} \cdot z_{ij} \leq l_{max}, \forall i \in M, \forall j \in C \tag{39}$$

Given the fixed decisions $\overline{x}_i$, $\overline{y}_{ij}$, $\overline{z}_{ij}$, $\overline{o}_{ik}$ the subproblem is formulate as follows:

$$\begin{aligned} &\min \sum_{s\in S} p_s \Big(\sum_{i\in M}\sum_{j\in D}\sum_{k\in K} c_r \cdot l_{ij} \cdot r^s_{ijk} \\ &+\sum_{i\in M}\sum_{j\in C}\sum_{k\in K}\sum_{t\in T} c_v \cdot l_{ij} \cdot q^s_{ijkt} + \sum_{i\in M}\sum_{j\in M}\sum_{k\in K}\sum_{t\in T_t} c_t \cdot l_{ij} \cdot g^s_{ijkt} \\ &+\sum_{i\in M}\sum_{k\in K}\sum_{t\in T} h_{ik} \cdot OH^s_{ikt} + b_k \cdot BO^s_{ikt} + \sum_{i\in M}\sum_{j\in C}\sum_{k\in K}\sum_{t\in T} c_k \cdot q^s_{ijkt}\Big) \end{aligned} \tag{40}$$

s.t.

$$o^s_{ik} - \overline{o}_{ik} \leq 0, \forall i \in M, \forall k \in K \tag{41}$$

$$I^s_{ikt} - o_{ik} \leq 0, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{42}$$

$$\sum_{j\in M} g^s_{ijkt} \leq (V^i_{max} / v_k) \cdot \overline{x}_i, \forall i \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{43}$$

$$\sum_{i\in M} g^s_{ijkt} \leq (V^j_{max} / v_k) \cdot \overline{x}_j, \forall j \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{44}$$

$$o_{ik} - I^s_{ikT} - r^s_{ijk} \leq V^i_{max} \cdot (1 - \overline{y}_{ij}),\ \forall i \in M, \forall j \in D, \forall k \in K, \forall s \in S \tag{45}$$

$$-o_{ik} + I^s_{ikT} + r^s_{ijk} \leq V^i_{max} \cdot (1 - \overline{y}_{ij}),\ \forall i \in M, \forall j \in D, \forall k \in K, \forall s \in S \tag{46}$$

$$I^s_{ik0} - o_{ik} = -\sum_{j\in C} d^s_{jk1} \cdot \overline{z}_{ij}, \forall i \in M, \forall k \in K, \forall s \in S \tag{47}$$

$$I^s_{ikt} - I^s_{ik(t-1)} - \sum_{j\in M} g^s_{ijkt} + \sum_{j\in M} g^s_{jikt} = -\sum_{j\in C} d^s_{jkt} \cdot \overline{z}_{ij}, \forall i \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{48}$$

$$I^s_{ikt} - I^s_{ik(t-1)} = -\sum_{j\in C} d^s_{jkt} \cdot \overline{z}_{ij}, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, t \notin T_t\ , \forall s \in S \tag{49}$$

$$OH^s_{ik1} - o_{ik} + \sum_{j\in C} q^s_{ijk1} = 0, \forall i \in M, \forall k \in K, \forall s \in S \tag{50}$$

$$OH^s_{ikt} - OH^s_{ik(t-1)} - \sum_{j\in M} g^s_{ijkt} + \sum_{j\in M} g^s_{jikt} + \sum_{j\in C} q^s_{ijkt} = 0, \forall i \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \tag{51}$$

$$OH^s_{ikt} - OH^s_{ik(t-1)} + \sum_{j\in C} q^s_{ijkt} = 0, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, t \notin T_l, \forall s \in S \quad (52)$$

$$BO^s_{ik1} + \sum_{j\in C} q^s_{ijk1} = \sum_{j\in C} d^s_{jk1} \cdot z_{ij}, \forall i \in M, \forall k \in K, \forall s \in S \quad (53)$$

$$BO^s_{ikt} - BO^s_{ik(t-1)} + \sum_{j\in C} q^s_{ijkt} = \sum_{j\in C} d^s_{jkt} \cdot \bar{z}_{ij}, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, \forall s \in S \quad (54)$$

$$\sum_{\tau\in T, \tau\le t} q^s_{ijk\tau} \le \sum_{\tau\in T, \tau\le t} d^s_{jk\tau} \cdot \bar{z}_{ij}, \forall i \in M, \forall j \in C, \forall k \in K, \forall t \in T, \forall s \in S \quad (55)$$

$$\sum_{j\in C} q^s_{ijkt} - OH^s_{ik(t-1)} \le 0, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, \forall s \in S \quad (56)$$

$$OH^s_{ikt} - I^s_{ikt} \ge 0, \forall i \in M, k \in K, t \in T, \forall s \in S \quad (57)$$

$$BO^s_{ikt} + I^s_{ikt} \ge 0, \forall i \in M, k \in K, t \in T, \forall s \in S \quad (58)$$

$$q^s_{ijkt} - q^l_{ijkt} = 0, \forall l \in S_{ts}, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T, \forall s \in S \quad (59)$$

$$g^s_{ijkt} - g^l_{ijkt} = 0, \forall l \in S_{ts}, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T, \forall s \in S \quad (60)$$

Appendix 2 Dual of the Benders Subproblem

By associating constraints (41), (42), …, (46) with dual variables $\pi^s_{1,ik}$, $\pi^s_{2,ikt}$, $\ldots \pi^s_{6,ijk}$ respectively; constraints (47), (48), (49) with dual variables $\pi^s_{7,ikt}$; constraints (50), (51), (52) with dual variables $\pi^s_{8,ikt}$; constraints (53) and (54) with dual variables $\pi^s_{9,ikt}$, and constraints (55), (56), …, (60) with dual variables $\pi^s_{10,ijkt}$, $\pi^s_{11,ikt}$, $\ldots \pi^{sl}_{15,ijkt}$ respectively, the dual of the subproblem can be formulated as follows:

$$\begin{aligned}
&\max \sum_{s\in S}\sum_{i\in M} \pi^s_{1,ik} \cdot V^i_{max} \cdot \bar{x}_i + \sum_{s\in S}\sum_{i\in M}\sum_{k\in K}\sum_{t\in T_t} \pi^s_{3,ikt} \cdot (V^i_{max} / v_k) \cdot \bar{x}_i \\
&+\sum_{s\in S}\sum_{j\in M}\sum_{k\in K}\sum_{t\in T_t} \pi^s_{4,jkt} \cdot (V^j_{max} / v_k) \cdot \bar{x}_j + \sum_{s\in S}\sum_{i\in M}\sum_{j\in D}\sum_{k\in K} \left(\pi^s_{5,ijk} + \pi^s_{6,ijk}\right) \cdot V^i_{max} \cdot \left(1 - \bar{y}_{ij}\right) \\
&+\sum_{s\in S}\sum_{i\in M}\sum_{k\in K}\sum_{t\in T} \left(-\pi^s_{7,ikt} + \pi^s_{9,ikt}\right) \cdot \sum_{j\in C} d^s_{jkt} \cdot \bar{z}_{ij} \\
&+\sum_{s\in S}\sum_{i\in M}\sum_{j\in C}\sum_{k\in K}\sum_{t\in T} \pi^s_{10,ijkt} \cdot \sum_{\tau\in T, \tau\le t} d^s_{jk\tau} \cdot \bar{z}_{ij}
\end{aligned} \quad (61)$$

s.t.

For o_{ik} :

$$v_k \cdot \pi^s_{1,ik} - \sum_{\tau\in T} \pi^s_{2,ik\tau} + \sum_{j\in D} \left(\pi^s_{5,ijk} - \pi^s_{6,ijk}\right) - \pi^s_{7,ik1} - \pi^s_{8,ik1} \le 0, \forall i \in M, \forall k \in K, \forall s \in S \quad (62)$$

For q^s_{ijkt} :

$$\pi^s_{8,ikt} + \pi^s_{9,ikt} + \sum_{\tau\in T, \tau\ge t} \pi^s_{10,ijk\tau} + \pi^{sl}_{14,ijkt} - \pi^{ls}_{14,ijkt} \le c_v \cdot l_{ij} + c_k, \forall i \in M, \forall j \in C, \forall k \in K, t = 1, \forall s \in S \quad (63)$$

$$\pi^s_{8,ikt} + \pi^s_{9,ikt} + \sum_{\tau\in T, \tau\ge t} \pi^s_{10,ijk\tau} + \pi^s_{11,ikt} + \pi^{sl}_{14,ijkt} - \pi^{ls}_{14,ijkt} \le c_v \cdot l_{ij} + c_k, \forall i \in M, \forall j \in C, \forall k \in K, \forall t \in T \setminus \{1\}, \forall s \in S \quad (64)$$

For g^s_{ijkt} :

$$\pi^s_{3,ikt} + \pi^s_{4,jkt} - \pi^s_{7,ikt} + \pi^s_{7,jkt} - \pi^s_{8,ikt} + \pi^s_{8,jkt} + \pi^{sl}_{15,ijkt} - \pi^{ls}_{15,ijkt} \le c_t \cdot l_{ij}, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T_t, \forall s \in S \quad (65)$$

For r^s_{ijk} :

$$-\pi^s_{5,ijk} + \pi^s_{6,ijk} \le c_r \cdot l_{ij}, \forall i \in M, \forall j \in D, \forall k \in K, \forall s \in S \quad (66)$$

For I^s_{ikt} :

$$\pi^s_{2,ikt} + \pi^s_{7,ikt} - \pi^s_{7,ik(t+1)} - \pi^s_{12,ikt} + \pi^s_{13,ikt} = 0, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{\mathrm{T}\}, \forall s \in S \quad (67)$$

$$\pi^s_{2,ikt} + \sum_{j\in D} \left(-\pi^s_{5,ijk} + \pi^s_{6,ijk}\right) + \pi^s_{7,ikt} - \pi^s_{12,ikt} + \pi^s_{13,ikt} = 0, \forall i \in M, \forall k \in K, t = \mathrm{T}, \forall s \in S \quad (68)$$

For OH_{ikt}^{s} :

$$\pi_{8,ikt}^{s} - \pi_{8,ik(t+1)}^{s} - \pi_{11,ik(t+1)}^{s} + \pi_{12,ikt}^{s} \le h_{ik}, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{\mathrm{T}\}, \forall s \in S \tag{69}$$

$$\pi_{8,ikt}^{s} + \pi_{12,ikt}^{s} \le h_{ik}, \forall i \in M, \forall k \in K, t = \mathrm{T}, \forall s \in S \tag{70}$$

For BO_{ikt}^{s} :

$$\pi_{9,ikt}^{s} - \pi_{9,ik(t+1)}^{s} + \pi_{13,ikt}^{s} \le b_{k}, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{\mathrm{T}\}, \forall s \in S \tag{71}$$

$$\pi_{9,ikt}^{s} + \pi_{13,ikt}^{s} \le b_{k}, \forall i \in M, \forall k \in K, t = \mathrm{T}, \forall s \in S \tag{72}$$

Where:

$$\pi_{1,ik}^{s} \le 0, \forall i \in M, \forall s \in S \tag{73}$$

$$\pi_{2,ikt}^{s} \le 0, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{74}$$

$$\pi_{3,ikt}^{s} \le 0, \forall i \in M, \forall k \in K, \forall t \in T_{t}, \forall s \in S \tag{75}$$

$$\pi_{4,jkt}^{s} \le 0, \forall j \in M, \forall k \in K, \forall t \in T_{t}, \forall s \in S \tag{76}$$

$$\pi_{5,ijk}^{s} \le 0, \forall i \in M, \forall j \in D, \forall k \in K, \forall s \in S \tag{77}$$

$$\pi_{6,ijk}^{s} \le 0, \forall i \in M, \forall j \in D, \forall k \in K, \forall s \in S \tag{78}$$

$$\pi_{7,ikt}^{s} \in R, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{79}$$

$$\pi_{8,ikt}^{s} \in R, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{80}$$

$$\pi_{9,ikt}^{s} \in R, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{81}$$

$$\pi_{10,ijkt}^{s} \le 0, \forall i \in M, \forall j \in C, \forall k \in K, \forall t \in T, \forall s \in S \tag{82}$$

$$\pi_{11,ikt}^{s} \le 0, \forall i \in M, \forall k \in K, \forall t \in T \setminus \{1\}, \forall s \in S \tag{83}$$

$$\pi_{12,ikt}^{s} \ge 0, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{84}$$

$$\pi_{13,ikt}^{s} \ge 0, \forall i \in M, \forall k \in K, \forall t \in T, \forall s \in S \tag{85}$$

$$\pi_{14,ijkt}^{sl} \in R, \forall i \in M, \forall j \in C, \forall k \in K, \forall t \in T, \forall s \in S, \forall l \in S_{ts} \tag{86}$$

$$\pi_{15,ijkt}^{sl} \in R, \forall i \in M, \forall j \in M, \forall k \in K, \forall t \in T, \forall s \in S, \forall l \in S_{ts} \tag{87}$$

References

[1] Schorung, M., 2024. Quick commerce and the evolving business models of the food retail industry - Investigating the quick commerce supply chain and the urban impacts of dark stores. Transportation Research Procedia 79, 305–312.

[2] Buldeo Rai, H., Mariquivoi, J., Schorung, M., Dablanc, L., 2023. Dark stores in the City of Light: Geographical and transportation impacts of 'quick commerce' in Paris. Research in Transportation Economics 100, 101333.

[3] Huang, M., Yen, B.P.C., 2021. Driving Forces for Digital Transformation – Case Studies of Q-Commerce.

[4] Ganapathy, V., Dr Chithambar Gupta, 2023. Critical success factors for quick commerce grocery delivery in India: an exploratory study. Sust. Agric. Food Env. Res. 12 (1).

[5] Yang, X., Ostermeier, M., Hübner, A., 2024. Winning the race to customers with micro-fulfillment centers: an approach for network planning in quick commerce. Cent Eur J Oper Res 32 (2), 295–334.

[6] Montecinos, J., Ouhimmou, M., Chauhan, S., Paquet, M., Gharbi, A., 2021. Transport carriers' cooperation on the last-mile delivery in urban areas. Transportation 48 (5), 2401–2431.

[7] Muñoz-Villamizar, A., Montoya-Torres, J.R., Vega-Mejía, C.A., 2015. Non-Collaborative versus Collaborative Last-Mile Delivery in Urban Systems with Stochastic Demands. Procedia CIRP 30, 263–268.

[8] Collins, B.C. Facility Location Optimization for Last-mile Delivery.

[9] Knežević, N., Milenković, M., Bojović, N. (Eds.), 2021. OPTIMAL LOCATION ANALYSIS OF MICRO DISTRIBUTION CENTERS FOR LAST MILE DELIVERY. University of Belgrade, Faculty of Transport and Traffic Engineering.

[10] Mrabti, N., Hamani, N., Boulaksil, Y., Amine Gargouri, M., Delahoche, L., 2022. A multi-objective optimization model for the problems of sustainable collaborative hub location and cost sharing. Transportation Research Part E: Logistics and Transportation Review 164, 102821.

[11] Novotná, M., Švadlenka, L., Jovčić, S., Simić, V., Cavallaro, F., 2022. Micro-hub location selection for sustainable last-mile delivery. PLoS ONE 17 (7), e0270926.

[12] Bok, M. de, Giasoumi, S., Tavasszy, L., Thoen, S., Nadi, A., Streng, J., 2024. A simulation study of the impacts of micro-hub scenarios for city logistics in Rotterdam. Research in Transportation Business & Management 56, 101186.

[13] Castillo, C., Panadero, J., Alvarez-Palau, E.J., Juan, A.A., 2024. Towards greener city logistics: an application of agile routing algorithms to optimize the distribution of micro-hubs in Barcelona. Eur. Transp. Res. Rev. 16 (1), 44.

[14] Huang, Z., Huang, W., Guo, F., 2020. Integrated sustainable planning of micro-hub network with mixed routing strategy. Computers & Industrial Engineering 149, 106872.

[15] Nieto-Isaza, S., Fontaine, P., Minner, S., 2022. The value of stochastic crowd resources and strategic location of mini-depots for last-mile delivery: A Benders decomposition approach. Transportation Research Part B: Methodological 157, 62–79.

[16] Ackva, C., Ulmer, M.W., 2024. Consistent routing for local same-day delivery via micro-hubs. OR Spectrum 46 (2), 375–409.

[17] Ji, H., Yang, S., Jia, B., Zhang, M., Su, B., 2024. Study on Location Selection of Urban Two-Level Joint Express Delivery Stations Considering Fair Cost Allocation among Enterprises. Transportation Research Record: Journal of the Transportation Research Board 2678 (10), 1551–1568.

[18] Yang, X., Ostermeier, M., Hübner, A., 2024. Winning the race to customers with micro-fulfillment centers: an approach for network planning in quick commerce. Cent Eur J Oper Res 32 (2), 295–334.

[19] Max Shen, Z.-J., Qi, L., 2007. Incorporating inventory and routing costs in strategic location models. European Journal of Operational Research 179 (2), 372–389.

[20] Ahmadi Javid, A., Azad, N., 2010. Incorporating location, routing and inventory decisions in supply chain network design. Transportation Research Part E: Logistics and Transportation Review 46 (5), 582–597.

[21] Zheng, X., Yin, M., Zhang, Y., 2019. Integrated optimization of location, inventory and routing in supply chain network design. Transportation Research Part B: Methodological 121, 1–20.

[22] Araya-Sassi, C., Paredes-Belmar, G., Gutiérrez-Jarpa, G., 2020. Multi-commodity inventory-location problem with two different review inventory control policies and modular stochastic capacity constraints. Computers & Industrial Engineering 143, 106410.

[23] Amiri-Aref, M., Klibi, W., Babai, M.Z., 2018. The multi-sourcing location inventory problem with stochastic demand. European Journal of Operational Research 266 (1), 72–87.

[24] Ahmadi, G., Torabi, S.A., Tavakkoli-Moghaddam, R., 2016. A bi-objective location-inventory model with capacitated transportation and lateral transshipments. International Journal of Production Research 54 (7), 2035–2056.

[25] Puttemans, I., an Caris, Braekers, K., 2025. The integration of location and inventory decisions in supply chain networks: A literature review and future prospects. Transportation Research Part E: Logistics and Transportation Review 195, 103970.

[26] Magnanti, T.L., Wong, R.T., 1981. Accelerating Benders Decomposition: Algorithmic Enhancement and Model Selection Criteria. Operations Research 29 (3), 464–484.

Biography

Zhu Zhu (*1999) is a PhD student at the Laboratory of Computer Science and Digital Society, University of Technology of Troyes (UTT), France. She obtained her master degree in industrial engineering from the School of Management, Xi'an Jiaotong University in 2024, specializing in logistics and supply chain management.

Haoxun Chen (*1964) is a full professor at the Laboratory of Computer Science and Digital Society, University of Technology of Troyes (UTT), France. He obtained his doctoral degree in systems engineering from the Systems Engineering Institute, Xi'an Jiaotong University in 1990. He has published more than 180 papers in international journals and conferences. His research interests include supply chain design and inventory management, collaborative logistics, production planning and scheduling, data envelopment analysis, and game theory.

8th Conference on Production Systems and Logistics

Analyzing Actor Interaction Gaps for the Digital Product Passport in a Multi-Stakeholder Ecosystem

Doğan Efe, Elena Andrushchenko, Roland Jochem
[1]*Chair of Quality Science, Technical University of Berlin, Berlin, Germany*

Abstract

The Digital Product Passport (DPP) is emerging as a key instrument in the European Union's transition toward a transparent, traceable, and circular economy. Building on the regulatory frameworks of the Battery Regulation, the Ecodesign for Sustainable Products Regulation (ESPR), and the draft standards of the DIN EN 18216 – 18246 series, the DPP establishes a service-oriented digital infrastructure for standardized information exchange across product life cycles. Its implementation requires a complex ecosystem of interacting stakeholders, including Economic Operators (EO), Service Providers (SP), and downstream actors each governed by distinct access and modification rights. Within this ecosystem, the management of granular data interactions poses major technical and organizational challenges. A multi-layered architecture based on Role-Based Access Control (RBAC) is essential to ensure data integrity, interoperability, and security across distributed systems. The interaction between actors and central DPP system functions introduces potential failure points, particularly where access permissions or sequencing logic are not clearly defined in the legislation. Moreover, uncertainties within the standardization process, combined with the limited timeframe for mandatory regulatory implementation, increase the complexity of establishing a compliant and reliable DPP infrastructure. In this context, the DPP ecosystem illustrates the extent to which regulatory requirements, technical system design, and stakeholder coordination must align to create a secure and interoperable foundation for product governance in Europe.

Keywords

Circular Economy; Digital Product Passport; Interoperability; Regulatory Compliance; Requirements Engineering

1. Introduction

The transition to a circular economy requires a profound restructuring of how products are manufactured, used, and recovered, which in turn demands granular, lifecycle spanning information about product characteristics, provenance, and environmental performance. The Digital Product Passport (DPP) has emerged as a key policy instrument to provide such information by establishing a digital record of product information maintained throughout the product's life, thereby facilitating reuse, repair, remanufacturing, and recycling while enabling regulatory oversight and informed decision making by value chain actors [1].

The regulatory foundation for this ecosystem is provided by two complementary instruments: the Battery Regulation, which introduces mandatory battery passports for specific battery categories from February 2027 and constitutes the first operational implementation of the DPP concept, and the ESPR, which establishes a cross-sectoral framework for extending digital passports to additional product groups through delegated acts [2], [3]. These regulations are further operationalized through a set of draft standards from the

DOI: https://doi.org/10.15488/21012
ISSN: 2701-6277

DIN EN 18216 – 18246 series, developed within the european standardization comittee CEN-CENELEC Joint Technical Committee (JTC 24). The series specifies, data exchange protocols, unique identifiers and data carriers, data storage requirements, application programming interfaces (APIs), interoperability mechanisms, access control, and data authentication. Together, these instruments define the structural building blocks of the DPP ecosystem while leaving considerable discretion regarding business logic, process orchestration, and governance mechanisms.

The DPP ecosystem is inherently multi-stakeholder in nature, encompassing Economic Operators (EO), Service Providers (SP), regulatory authorities, notified bodies, repair and maintenance providers, refurbishers, recyclers, and consumers [4]. Each of these actor categories has distinct responsibilities, data needs, and rights, producing an authorization landscape that must balance the objectives of transparency, circularity, and regulatory control with the protection of commercially sensitive or privacy-relevant information. Role-based access control (RBAC), lifecycle-dependent permission management, and verifiable digital identities become central design elements to manage these tensions in a scalable and auditable manner.

At the same time, the DPP does not operate as a static repository but as a living, distributed infrastructure in which actor interactions occur through sequences of API mediated operations that may involve multiple independent systems, asynchronous communication, and overlapping process executions across organizational boundaries [5]. While existing standards describe the available operations and roles at a structural level, they are largely agnostic regarding the orchestration rules that govern how these elements should interact under realistic conditions.

This gap motivates the central research question addressed in this paper: How must actor interaction models for the DPP be structured to ensure governance-stable, orchestration-compliant, and interoperable system operation across a heterogeneous multi-stakeholder environment? While prior work has addressed DPP data governance [6] and ecosystem value creation [7], the systematic analysis of interaction gaps and their orchestration implications remains insufficiently addressed in the literature [8].

Against this background, this paper pursues two interrelated objectives. First, it synthesizes the ecosystem-level architectural and regulatory perspective on DPPs with a detailed analysis of actor interactions, highlighting the interplay between interoperability and RBAC. Second, it identifies critical gaps between high-level standards and concrete operational behavior, and proposes a solution for a more explicit orchestration framework capable of ensuring consistent and trustworthy DPP operation in a multi-stakeholder environment.

While this paper focuses on a multi-layered architecture based on RBAC and standardized API interactions as defined in the DIN EN 18216–18246 series, several complementary digital modelling approaches for DPP implementation exist and warrant acknowledgement. The Asset Administration Shell (AAS) provides a semantic framework for Industry 4.0 components and has been proposed both as an interoperability enabler for cross-industry data exchange [9] and as an implementation substrate for DPP functionality within Gaia-X-based data spaces [10], including for advanced Maintenance, Repair, and Overhaul (MRO) processes integrating proactive AAS with DPP lifecycle management [11]. Furthermore, decentralized Digital Twin models leveraging blockchain or Distributed Ledger Technology have been proposed to ensure data integrity, immutability, and distributed trust [12], with ongoing work exploring how such architectures can be applied to battery and product lifecycle traceability at scale [13], [14]. For actor authentication and authorization, decentralized identity frameworks based on W3C Decentralized Identifiers and Verifiable Credentials offer an alternative to centralized identity providers, enabling tamper-proof credential verification without reliance on a single trusted authority [15]. These approaches offer different trade-offs between centralized coordination and fully distributed data sovereignty. The DPP as specified under the ESPR and the Battery Regulation, however, constitutes a regulation-driven, legally binding paradigm that mandates a specific governance structure, differentiated access tier model, and lifecycle management

obligation – motivating the focus of this paper on the interaction gaps arising specifically within this regulatory framework.

2. DPP Ecosystem Architecture and Interoperability

2.1 Distributed System Architecture

The DPP ecosystem is designed as a distributed and largely decentralized system in which passport data are hosted under the responsibility of the relevant EO, rather than in a single central database [12]. In this paradigm, the EO placing a product on the market remains the passport custodian and may either host the passport within its own IT infrastructure while relying on a designated backup SP, or delegate hosting and operational responsibility to a main SP that, in turn, coordinates with a backup SP to ensure continuity and resilience.

A purely decentralized model, however, would make it difficult to discover, verify, and monitor the existence of passports across the internal market. To address this limitation, the ESPR mandates the establishment of a central EU registry that records unique identifiers for products placed on the market and links these identifiers to the locations of their associated passports [2]. The registry functions as a directory rather than as the primary data store, enabling authorities and other actors to locate the responsible custodian service while maintaining the distributed hosting approach [5]. This combination results in a hybrid architecture that integrates decentralized data storage with centralized coordination and oversight functions.

Finally, each physical product must be reliably linked to its digital passport via globally unique identifiers and standardized data carriers, such as QR-codes or Radio Frequency Identification (RFID) tags, affixed to the product or its packaging. These carriers encode a resolvable identifier that allows any actor to locate the corresponding passport through the registry and custodian services, thereby closing the physical–digital loop essential for practical use in manufacturing, maintenance, and recycling contexts [16], [17].

2.2 Interoperability Standards and Information Models

Interoperability constitutes a foundational requirement for the DPP ecosystem, given the diversity of stakeholders, IT systems, and product categories involved. The ESPR explicitly requires that passport data be based on open standards, be provided in a structured, machine-readable, and searchable format, and be transferable via an interoperable, vendor-neutral data exchange network, thereby mandating technical, semantic, and organizational interoperability. To operationalize these requirements, draft European standards (DIN EN 18216–18223) define common data formats, serialization mechanisms, and interface specifications for interacting with passport data [16], [17], [18].

2.3 Stakeholder Landscape and Data Flows

EOs are the primary actors with direct legal obligations for DPP creation and maintenance. This category comprises six subgroups with distinct duties. Manufacturers are actors who produce or have products manufactured and market them under their own name, bearing primary responsibility for creating DPPs when placing products on the market. Importers place third-country products on the Union market and assume responsibility for DPP compliance. Distributors are supply chain actors other than manufacturers or importers who make products available without being responsible for DPP content creation. Dealers are distributors offering products to end users, including via distance selling [4].

Other value chain actors require DPP access but do not bear primary creation responsibility. Customers may want to access controlled DPP data, especially data derived from product usage. Professional repairers require technical specifications, repair instructions, and component information. Recyclers require material composition and disassembly data. Market surveillance authorities need broad access for compliance verification, while customs authorities verify that imported products have valid DPPs [4].

SPs form the third actor category, offering technical infrastructure for DPP data management. Main DPP SPs are authorized to process and make DPP data available to actors with access rights. Backup SPs host redundant copies to ensure availability and maintain historical versions to support long-term traceability and compliance [4], [16].

This taxonomic framework reveals the fundamental characteristic of the DPP ecosystem: multiple actor categories with divergent objectives and access requirements must interact with shared data resources. EOs seek to minimize disclosure of commercially sensitive information while meeting mandatory transparency obligations. Regulatory authorities require comprehensive access to verify compliance. Professional repairers and recyclers require detailed technical data to perform their functions effectively. Customers increasingly demand product sustainability information. Balancing these competing requirements through technical access control mechanisms represents a core challenge in DPP system design.

3. Role-Based Access Control and Security

Access to DPP information cannot be modeled as a binary public-versus-private distinction; instead, it must reflect a nuanced spectrum of data sensitivity and stakeholder roles. The regulatory framework recognizes this complexity by mandating differentiated access to passport data depending on data type and stakeholder typology, implying that each data element must be associated with an access category and that stakeholders must be authenticated and authorized accordingly [19]. The Battery Regulation provides a concrete implementation example by defining at least three access tiers: publicly accessible information; information restricted to regulatory authorities and the European Commission; and information accessible only to third parties demonstrating a legitimate interest, such as professional repairers or recyclers [3].

This multi-tiered approach allows sensitive or commercially confidential information such as proprietary process parameters to be shielded from general public access while remaining available to actors who require it for safety, compliance, or circularity functions. At the same time, a minimum level of openness is preserved for citizen-facing purposes, such as informing consumers about environmental performance, reparability, or warranty conditions. Designing and enforcing these distinctions requires not only regulatory definitions but also technical mechanisms within DPP platforms to filter responses, anonymize or aggregate data when necessary, and log access events for auditing purposes [9], [6].

Because access entitlements evolve as products move through lifecycle stages and responsibilities shift between actors, these lifecycle transitions necessitate timely revocation, reassignment, and creation of permissions so that former custodians lose write access when their responsibility ends, new custodians gain appropriate rights, and authorization gaps or overlaps are avoided. Consequently, DPP platforms must implement mechanisms for rule-based, event-driven updates to authorization policies, triggered by lifecycle status changes, ownership transfers, or regulatory events. RBAC in a distributed, cross-border ecosystem presupposes reliable digital identification and authentication of actors. The question of who determines and verifies that an entity is indeed a certified recycler, a notified body, or a market surveillance authority is central to security and trust. Emerging approaches advocate the use of digital identity frameworks, to provide cryptographically verifiable evidence of an actor's role and qualifications [17], [20], [21].

In such a model, a trusted authority such as a national regulator or accredited certification body issues digital credentials attesting to an organization's role (e.g., "licensed recycler") When interacting with a DPP system, the actor presents these credentials, which are then verified cryptographically and used to determine access rights to specific data elements or API operations [2], [22]. Beyond access control, the DPP ecosystem must uphold key security properties, including integrity, authenticity, non-repudiation, and availability. Integrity requires that passport data must be protected against unauthorized modification and that any change must be reliably detectable; authenticity demands assurance that data originate from legitimate sources;

non-repudiation ensures that actors cannot deny their actions; and availability mandates that data remain accessible under appropriate conditions throughout the product lifecycle.

4. Actor Interactions and Process Orchestration

4.1 CRUD Operations in a Distributed DPP Context

At a technical level, the DPP ecosystem consists of operations that resemble the standard CRUD (Create, Read, Update, Delete) pattern known from information systems. Read operations are conceptually simpler because they do not change the system state and generally involve a single requesting actor and the custodian service hosting the passport. Nonetheless, read operations must enforce access control policies by filtering or tailoring the returned datasets based on the requester's role and credentials, ensuring that sensitive fields are excluded or anonymized where necessary [4]. Write operations, including update, and delete, are inherently more complex because they change the system state and often involve multiple subsystems such as custodian platforms, backup services, and the central registry in a coordinated sequence.

4.2 Illustrative Scenarios for Role-Dependent Access Control

4.2.1 Reading public DPP data

Reading publically available DPP data will be a standard process after the implementation of the DPP in 2027. All actors within the DPP ecosystem can access specific Battery Passport data of public interest using a QR-code or RFID tag affixed to the product or included within the product documentation. Access to publicly available data shall be enabled via a mobile device capable of scanning the QR-code (e.g., a smartphone) or reading the RFID tag. Upon scanning the identifier, the user shall be directed to a web-based interface where the publicly accessible Battery Passport data is rendered in a structured, human-readable format. The solution shall ensure seamless redirection from the physical identifier to the corresponding online resource and shall provide public accessibility without authentication barriers for data classified as public [17]. To complement the description above, Table 1 provides a structured overview of the actors and the corresponding access mechanism for public DPP data.

Table 1: Overview of Actors and Access Preconditions for Public Battery Passport Data

Actor	Role	Responsibilities	Auth. Protocol
EO	DPP Custodian / Website Host	Publishes and maintains publicly accessible Battery Passport data. Ensures that the QR-code or RFID resolves to a persistent URL. Hosts the web interface rendering public DPP data in human-readable format. Ensures data accuracy, availability, and separation between public and restricted data domains.	—
End User (e.g. Consumer)	Public Data Viewer	Scans QR-code using a mobile device. Accesses the redirected web interface. Views publicly available DPP data.	No authentication required

4.2.2 Maintaining DPP data

Updating DPP data will constitute a standard process once DPPs are operationalized. In contrast to the retrieval of publicly available information, updates are restricted to authorized actors and shall be performed through authenticated, API-mediated interactions with the passport custodian. Depending on the lifecycle stage and the applicable regulatory requirements, actors such as EOs, SPs, professional repairers, and recyclers may need to submit updates for example to record maintenance actions, ownership changes, compliance information, or end-of-life status transitions. Access to update functionalities shall therefore be enabled through a secure authentication mechanism (e.g., OAuth 2.0 / OpenID Connect), ensuring that the requesting actor's identity and role are verified and that only permitted data elements and operations can be

modified [5] [18]. Table 2 provides a structured overview of the actors and the corresponding access mechanisms for updating Battery Passport data.

Table 2: Overview of Actors and Access Preconditions for Maintaining DPP Data

Actor	Role	Responsibilities	Auth Protocol
EO	DPP Custodian	Validates identity, checks access rights, performs update, creates new DPP ID, archives previous DPP, initiates backup of DPP	eIDAS or oAuth 2.0
e.g., Recycler	Actor performing Update	Sends PATCH request with DPP ID and update payload. Must hold write rights for all requested attributes	eIDAS or oAuth 2.0
Backup SP	Storage Provider	Authenticates EO, creates new backup copy of updated DPP, archives previous backup	oAuth 2.0

Figure 1 provides an illustrative example of the ordinary sequence of the update process as derived from the current draft standards. Because the standards do not fully specify backup and archiving behavior, the sequence shown includes explicit assumptions regarding versioning, archival state transitions, and the creation of backup copies as well as the handling of IDs. The figure depicts a four-phase interaction between a third-party actor, the EO as passport custodian, and a backup SP:

A third-party actor initiates an authentication request to the EO using an OAuth 2.0 / OpenID Connect–based flow [18]. Upon successful authentication, the EO issues an authorization token that encodes the actor's role-based access rights, thereby enabling subsequent operations. The third-party actor submits a partial update to an existing passport via a PATCH request using the specific API endpoint. The EO applies the update and performs internal versioning/archiving by creating a new DPP version and assigning a new DPP identifier, while transitioning the previous version to an archived state and marking the new version as active. To establish a backup, the EO authenticates with the backup SP (again via OAuth 2.0), which returns an access token confirming authorization for the backup operation. The EO creates the backup by issuing a POST request to the specific API endpoint, transmitting the updated DPP content together with the newly assigned DPP identifier. The backup SP archives any previous backup copy and activates the new backup instance.

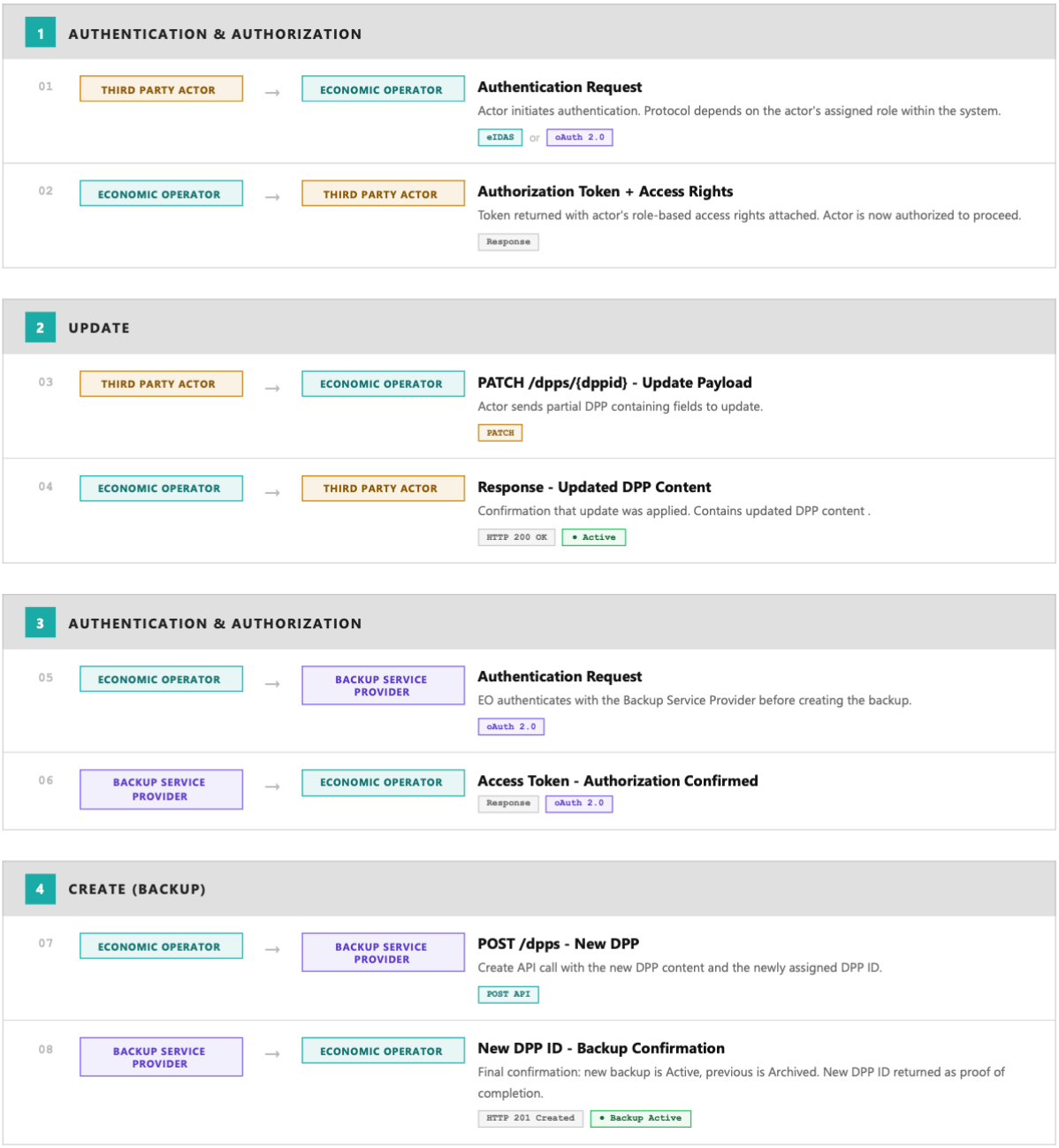


Figure 1: Ordinary Sequence for Updating DPP data

4.3 Challenges and Governance Gaps

Although the current draft standards allow the reading of public data, updating and backup sequence to be specified at a conceptual level, it remains unclear how these procedures will operate under real-world conditions.

A central gap concerns downstream actors such as recyclers. When a recycler receives a battery, the current standards do not clearly specify how controlled data elements can be accessed in order to submit an authorized update. The recycler may need to update the battery's lifecycle status to indicate that it has been recycled and is ready for disposal. This might be necessary as the EO and other authorized actors along the value chain should know about the final state of the battery.

In practice, the recycler will typically scan the battery's QR-code as shown in the example above, which redirects to a web page hosted by the EO and exposes only publicly available passport information. At this point, the standards provide no guidance on how the recycler can discover, request, and obtain access to

restricted data via the interface. The provided ordinary sequence in chapter 4.2.2 only works under the assumption that the recycler is already registered as such with the EO and has therefore access via the designated APIs.

In addition, this assumption is difficult to sustain at ecosystem scale. The DPP framework anticipates a broad and heterogeneous landscape of recyclers and refurbishers who may require write access to controlled data for legitimate lifecycle updates. In practice, EOs will likely collaborate primarily with known partners under contractual arrangements; however, a governance model in which access must be granted on a one-to-one basis to potentially large numbers of third-party SPs is neither operationally feasible nor aligned with the intended openness of downstream participation. The resulting mismatch between legally forseen access rights and permission administration constitutes a core actor-interaction gap: without a scalable mechanism for recognizing qualified third parties and assigning rights accordingly, controlled-data access for updates becomes ad hoc and illogical from an operational perspective.

One plausible solution would be to display, within the public view, an identifier that enables access to the DPP through the EU registry. However, registry-based access alone does not establish authentication and authorization by the EO, and the registry's expected search and discovery functionalities are not yet sufficiently specified in terms of scope or interfaces. As a result, the draft standards do not yet define a mechanism by which scanning the code can lead under appropriate authorization to controlled data access and subsequent update operations.

To address this gap, the approach illustrated in Figure 2 is proposed. After scanning the QR-code and being redirected to the EO's page, the recycler authenticates via a login button provided on the hosted webpage. The EO must represent and enforce the recycler's access rights on its own server. These rights derive from statutory requirements and therefore need to be encoded within the EO's authorization logic. However, to apply these rights correctly, the EO requires a reliable proof that the requesting party is in fact a recycler.

One approach is to rely on eIDAS-based digital identity mechanisms to establish the requester's identity and role. In this model, the EO obtains verifiable credentials or attributes that attest both to the person's identity and to a role/status such as "recycler." Based on these validated credentials, the EO can correctly map and grant the legally defined access rights to controlled DPP data.

The recycler is then granted role-appropriate access to controlled data through a dashboard, enabling the user to view restricted information and, where permitted, submit updates consistent with their assigned permissions as a recycler.

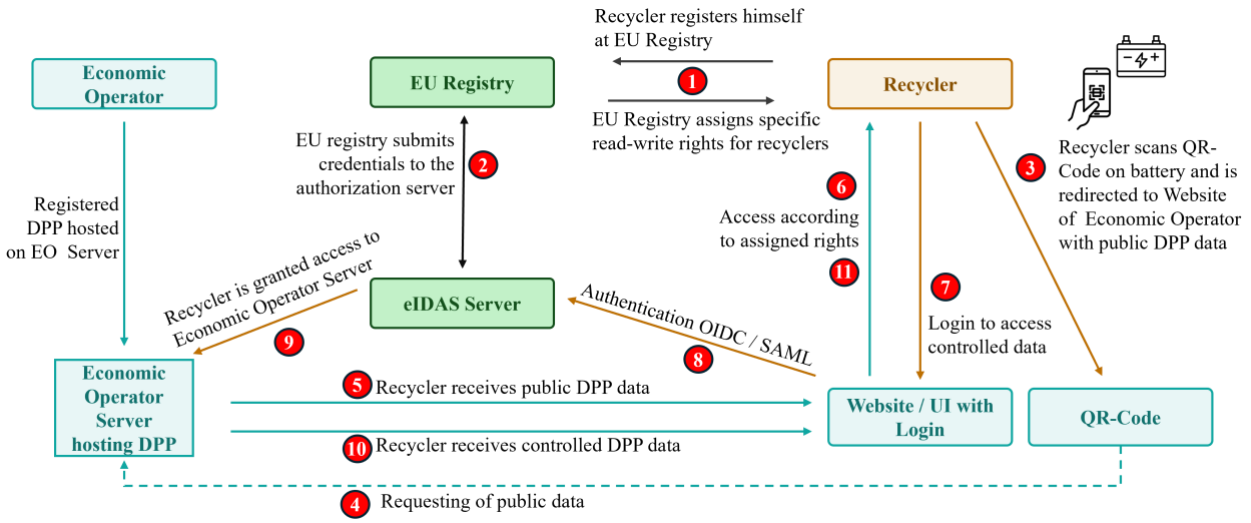


Figure 1: Recycler access to controlled DPP data

This example shows that despite the evident importance of sequencing, current standards primarily define individual operations and roles without prescribing detailed workflows, leaving the specification and enforcement of sequencing logic to implementers and sector-specific guidance. This flexibility supports adaptation to diverse product types but also increases the risk of inconsistent or insufficiently secured process designs across the ecosystem.

5. Discussion: Governance Needs Beyond Data Models and Interfaces

The existing regulatory and standardization framework provides extensive detail on data models, interfaces, and high-level role categories but deliberately stops short of prescribing concrete business logic for actor interactions. This design choice reflects the diversity of product groups and sectoral requirements, acknowledging that a single process blueprint is unlikely to suit all cases. However, this flexibility creates a gap between standardized capabilities and the actual behaviors that determine system robustness and security in practice.

The absence of harmonized orchestration patterns increases the risk that implementations will diverge in their handling of edge cases and failure conditions, potentially leading to interoperability issues, security weaknesses, or inconsistent user experiences across sectors and jurisdictions.

While regulations recognize the need to restrict access to certain data categories to actors with legitimate interests and specific qualifications, the concrete mechanisms for actor verification and credential issuance remain under development. The lack of a fully established, interoperable trust framework for DPP actors poses risks in two directions: unauthorized entities may misrepresent themselves to gain access to sensitive data, and legitimate actors may be unable to authenticate their status across different DPP platforms [4], [19].

To close this gap, additional governance mechanisms may be necessary, such as reference process models or implementation guidelines that specify recommended sequences and controls for common operations, as well as automated checks that detect anomalies such as missing passports, inconsistent status transitions, or unexplained inactivity. The development of such orchestration frameworks will likely require collaboration between regulators, standardization bodies, industry consortia, and technical communities to ensure both regulatory compliance and technical feasibility.

6. Conclusion

The investigation demonstrates that the Digital Product Passport represents more than a data container or a technical integration layer. It is an infrastructure whose reliability, trustworthiness, and regulatory effectiveness depend on aligned interaction between multiple independent actors operating within a distributed system architecture.

The analysis highlights that the current draft standards successfully establish fundamental building blocks, such as role models, access rights frameworks, and API-based connectivity. However, these elements do not yet ensure that interactions are inherently coherent, conflict-resilient, or governance-stable. A secure and robust DPP system requires more than declarative role definitions or standardized technical interfaces. It requires an explicitly designed orchestration of interactions across the ecosystem.

Such orchestration must address at least three key needs. It must define transparent and enforceable sequencing logic for multi-party processes. It must incorporate reliable mechanisms for validating and supervising roles with regulatory relevance. Finally, it must provide safeguards against concurrency conflicts, incomplete process execution, and inconsistent lifecycle states.

Future research should address several open challenges identified by this analysis. First, the scalability and cross-border interoperability of eIDAS-based digital identity frameworks for validating diverse actor roles, such as independent recyclers across different EU member states, requires further empirical and technical

investigation, particularly regarding dependable guarantees for data sovereignty and resilience against insider threats in federated architectures [23]. Second, formal process verification methods should be adapted to capture the sequencing and role-dependency constraints of DPP actor interactions, while simultaneously exploring the integration of automated anomaly detection and conflict resolution mechanisms to enhance system robustness against inconsistent lifecycle state transitions [24]. Third, investigating the potential of semantic technologies and knowledge graphs to bridge sectoral differences in data models, combined with the formalization of DPP-relevant regulatory requirements in machine-readable formats, remains a critical area for ensuring a truly universal and compliance-verifiable DPP infrastructure [25], [26].

Consequently, the long-term viability of the DPP cannot rely solely on individual actors' implementations or voluntary coordination. The ecosystem requires architectural governance that ensures predictable interaction behavior, harmonized process execution, and verifiable legitimacy of participating roles. Only under such conditions can the DPP fulfill its intended purpose as a secure, interoperable, and trustworthy foundation for product transparency, regulatory enforcement, and circular economy governance in Europe.

Acknowledgements

The project *BatteryPass-Ready* (funding period 01 April 2025 – 31 March 2027) is funded by the German Federal Ministry for Economic Affairs and Energy (BMWE) based on a resolution of the German Bundestag under grant agreement No. 16BZF363C.

References

[1] N. Fares, M.S. Islam, S.K. Jauhar, B. Küçükaltan, Towards an International Digital Product Passport: The New Paradigm of a Worldwide Circular Economy, Circ. Econ. Sustain. (2025). doi:10.1007/s43615-025-00690-5.

[2] European Union, Regulation (EU) 2024/1781 of the European Parliament and of the Council of 13 June 2024 establishing a framework for the setting of ecodesign requirements for sustainable products, Off. J. Eur. Union L (2024).

[3] European Union, Regulation (EU) 2023/1542 of the European Parliament and of the Council of 12 July 2023 concerning batteries and waste batteries, Off. J. Eur. Union L 191 (2023).

[4] DIN EN 18239 (prEN 18239:2025), Digitaler Produktpass - Management der Benutzerrechte, IT-Sicherheit und Geschäftsgeheimnisse [Digital Product Passport - Access rights management, information system security, and business confidentiality], 2025.

[5] DIN EN 18222 (prEN 18222:2025), Digitaler Produktpass - Programmierschnittstellen (APIs) für das Lebenszyklusmanagement und die Durchsuchbarkeit vom Produktpass [Digital Product Passport - Application Programming Interfaces (APIs) for the product passport lifecycle management and searchability], 2025.

[6] C. Ducuing, R.H. Reich, Data governance: Digital product passports as a case study, Compet. Regul. Netw. Ind. 24 (1) (2023) 3. doi:10.1177/17835917231152799.

[7] A. Gieß, F. Möller, Exploring the value ecosystem of digital product passports, J. Ind. Ecol. 29 (2) (2025) 561–573. doi:10.1111/jiec.13621.

[8] S.F. Jensen, J.H. Kristensen, A. Christensen, B.V. Wæhrens, An ecosystem orchestration framework for the design of digital product passports in a circular economy, Bus. Strategy Environ. 33 (7) (2024) 7100–7117. doi:10.1002/bse.3868.

[9] M. Kuhn, M. Baumann, F. Volz, L. Stojanović, Digital Product Passport Design Supporting the Circular Economy Based on the Asset Administration Shell, Sustainability 17 (3) (2025) 969. doi:10.3390/su17030969.

[10] K. Gleich, S. Behrendt, M. Hörger, M. Benfer, G. Lanza, An Asset Administration Shell-Based Digital Product Passport as a Gaia-X Service, Procedia CIRP 127 (2024) 224–229. doi:10.1016/j.procir.2024.07.039.

[11] M. Weiss, A. Kamtsiuris, F. Raddatz, G. Wende, Advanced MRO Processes in Industry 4.0 with Proactive Asset Administration Shell and Digital Product Passport, Int. J. Progn. Health Manag. 16 (2) (2025). doi:10.36001/ijphm.2025.v16i2.4508.

[12] R. Kannappan, J. Hatin, E. Bertin, N. Crespi, Enhancing Digital Product Passport Through Decentralized Digital Twins, in: NTMS 2025 - 11th IFIP Int. Conf. New Technol. Mobil. Secur., 2025: p. 300. doi:10.1109/ntms65597.2025.11076730.

[13] J. Mügge, J. Grosse Erdmann, T. Riedelsheimer, M.M. Manoury, S.O. Smolka, S. Wichmann, K. Lindow, Digital Twins within the Circular Economy: Literature Review and Concept Presentation, Sustainability 16 (7) (2024) 2748. doi:10.3390/su16072748.

[14] Z. Zhang, J. Gu, R. Li, W. Lian, C. Ma, J. Tian, Bridging digital and green transitions: A literature review of digital product passport systems, Renew. Sustain. Energy Rev. 228 (2026) 116600. doi:10.1016/j.rser.2025.116600.

[15] C. Mazzocca, A. Acar, S. Uluagac, R. Montanari, P. Bellavista, M. Conti, A Survey on Decentralized Identifiers and Verifiable Credentials, IEEE Commun. Surv. Tutor. (2025). doi:10.1109/COMST.2025.3543197.

[16] DIN EN 18223 (prEN 18223:2025), Digitaler Produktpass - System-Interoperabilität [Digital Product Passport - System interoperability], 2025.

[17] DIN EN 18219 (prEN 18219:2025), Digitaler Produktpass - Eindeutige Kennungen [Digital product passport - Unique identifiers], 2025.

[18] DIN EN 18216 (prEN 18216:2025), Digitaler Produktpass - Protokolle zum Datenaustausch [Digital product passport - Data exchange protocols], 2025.

[19] DIN EN 18246 (prEN 18246:2025), Digitaler Produktpass - Datenauthentifizierung, Zuverlässigkeit und Integrität [Digital product passport - Data authentication, reliability and integrity], 2025.

[20] R. Sandhu, E.J. Coyne, H.L. Feinstein, C.E. Youman, Role-based access control models, Computer 29 (2) (1996) 38. doi:10.1109/2.485845.

[21] I.I. García, F.D. Muñoz-Escoí, J.A. Aroca, F.J.F. Peñuela, Digital Product Passport Management with Decentralised Identifiers and Verifiable Credentials, arXiv preprint arXiv:2410.15758 (2024). doi:10.48550/arxiv.2410.15758.

[22] I. Plaza-Ortiz, A. Munoz-Arcentales, J. Salvachúa, C. Aparicio, G. Huecas, E. Barra, Authentication and authorization in Data Spaces: A relationship-based access control approach for policy specification based on ODRL, arXiv preprint arXiv:2505.24742 (2025). doi:10.48550/ARXIV.2505.24742.

[23] J. Lohmöller, J. Pennekamp, R. Matzutt, C.V. Schneider, E. Vlad, C. Trautwein, K. Wehrle, The unresolved need for dependable guarantees on security, sovereignty, and trust in data ecosystems, Data Knowl. Eng. 151 (2024) 102301. doi:10.1016/j.datak.2024.102301.

[24] T. Lopes, S. Guerreiro, Assessing business process models: a literature review on techniques for BPMN testing and formal verification, Bus. Process Manag. J. 29 (8) (2023) 133–162. doi:10.1108/BPMJ-11-2022-0557.

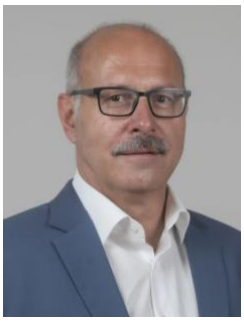
[25] B. Gallina, G.L. Steierhoffer, T.Y. Olesen, E. Parajdi, M. Aarup, Towards an ontology for process compliance with the (machinery) legislations, J. Softw. Evol. Process 37 (2025) e2728. doi:10.1002/smr.2728.

[26] T. Butler, D. Gozman, K. Lyytinen, The regulation of and through information technology: Towards a conceptual ontology for IS research, J. Inf. Technol. 38 (2) (2023) 86–107. doi:10.1177/02683962231181147.

Biography

Doğan Efe (*1995) studied mechanical engineering at the Technical University of Berlin. Following his time as a student assistant (2023 – 2024), Doğan Efe has been a Research Associate at the chair of Quality Science at TU Berlin's Institute for Machine Tools and Factory Management.

Elena Andrushchenko (*1988) studied industrial engineering at the Technical University of Berlin. Following her time as a student assistant (2023 – 2024), Elena Andrushchenko has been a Research Associate at the chair of Quality Science at TU Berlin's Institute for Machine Tools and Factory Management.



Prof. Dr.-Ing. Roland Jochem (*1962) studied Mechanical Engineering and earned his doctorate at TU Berlin. Following management roles at Fraunhofer IPK and Bosch-Siemens, and a professorship in Kassel, he led the Chair of Quality Science at Technical University of Berlin (2010 – 2025). He also serves as an Extraordinary Professor at Stellenbosch University (South Africa).

8th Conference on Production Systems and Logistics

Requirements Analysis For Near-Ground Object Detection Of Autonomous Mobile Robots

Thorge Lackner[1], Julian Hermann[1], Christian Kuhn[1], Daniel Palm[1,2]
[1]*Faculty of Sustainability and Technology - NXT, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen*
[2]*Fraunhofer Institute of Manufacturing Engineering and Automation, Alteburgstraße 150, 72762 Reutlingen*

Abstract

Autonomous Mobile Robots (AMR) increasingly operate in complex industrial environments, where reliable detection and classification of near-ground objects is critical for safe and efficient navigation. Near-ground obstacles such as cables, cables ducts, door thresholds, pallets and packaging materials pose significant challenges due to their low height, visual ambiguity and sensitivity to lighting, shadows and floor conditions. Existing perception systems often overlook these objects and fail under dynamic, industrial conditions, revealing an existing gap in current AMR perception capabilities. Therefore, this paper presents a structured requirements analysis for near-ground object detection in AMR. By integrating operational constraints, environmental influences and robot motion characteristics, this paper identifies the technical and functional requirements necessary for robust perception for near-ground objects at floor level. The resulting requirement framework provides a foundation for researching and developing improved detection and classification methods, enabling more reliable navigation in real-world industrial settings.

Keywords

Autonomous Mobile Robots; Object Classification; Near-Ground Objects; Artificial Intelligence; Computer Vision;

1. Introduction

Autonomous Mobile Robots (AMR) have become a critical pillar of intralogistics, enabling flexible material flows, dynamic routing, and a higher degree of autonomy [1–3]. However, despite remarkable advances in perception and navigation technologies over the past decade, AMR continue to face challenges in seemingly trivial situations within real-world factory environments [4,5,3]. One operationally relevant problem is the interaction with near-ground objects – such as door thresholds, ramps, cables, cable ducts, pallets, tools, or packaging residues – that are commonly found on industrial floors [6,5,7–9]. Although these objects are near-ground and often small in size, they can cause significant operational disruptions to AMR in case of collision, including emergency stops, navigation errors, and mechanical damage, creating a performance and safety gap between robot intelligence and the complexity of real-world environments [8,10].

The scientific relevance becomes evident when considering the characteristics of both the objects and the industrial environment. Industrial objects have highly varied geometries and textures, and limited height profiles, making them difficult to detect and, ultimately, classify with conventional laser scanners [8,9]. Additionally, objects' appearance is influenced by environmental factors such as shadows, glare, occlusions, and reflections [9]. At the same time, dynamic ad-hoc movements of AMR, introduce additional constraints: when AMRs accelerate, decelerate, experience driving vibrations or navigate uneven surfaces, perception systems must be robust and capable of real-time object detection [9]. False detections impact safety and

DOI: https://doi.org/10.15488/21013
ISSN: 2701-6277

operational efficiency, as near-ground obstacles can alter AMR's path, triggering collisions and manual restarts [8]. Even with 3D sensors and a generated point cloud, these challenges are exacerbated by the fact that most AMR intentionally filter out the area below the laser scanner plane resulting in "blind spots" for the near-ground objects [11,8,12]. The underlying problem is visualized in Figure 1.

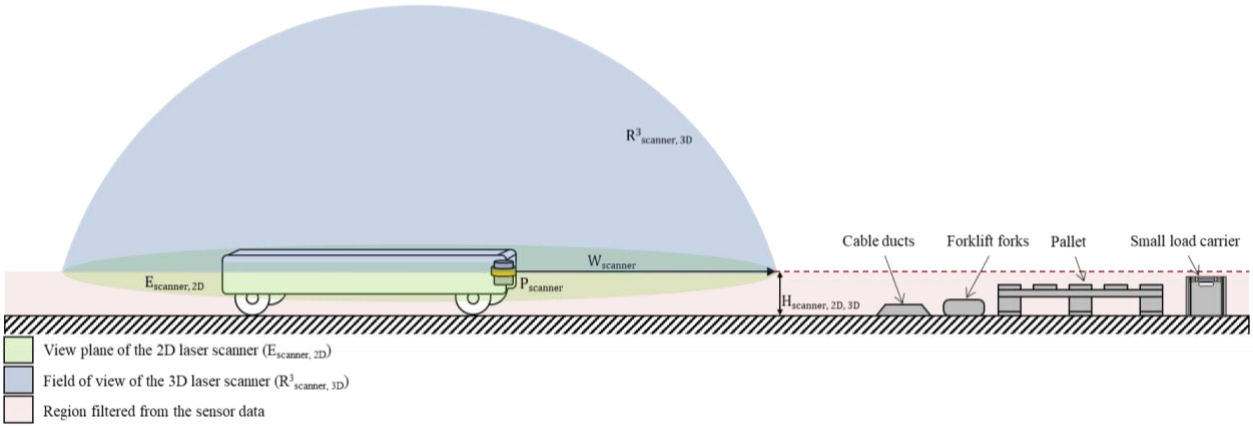


Figure 1: Fields of view and vision restrictions of AMR.

Against this background, the aim of this paper is to derive requirements for near-ground object detection of AMR. The analysis synthesises constraints from environmental, technical and operational factors to establish a requirement foundation. The resulting requirements enable a systematic development towards perception methods targeting near-ground objects on factory floors. To achieve this, the paper is organized as follows: Section one introduces the research. Section two outlines the background on near-ground perception of AMR. Section three presents the methodology. Section four consolidates the requirements while Section five discusses the results. Finally, Section six concludes the paper and identifies future research activities.

2. Background of near-ground perception of autonomous mobile robots

Despite using multiple sensors such as 2D and 3D Light Detection and Ranging (LiDAR) sensors, RGB cameras and depth cameras, robust perception of near-ground and small objects represents a critical but persistently underdeveloped capability in the field of AMR [7,9].

2D laser scanners (2D-LiDAR) remain commonly used sensors in industrial robotics due to their robustness, reliability, and measurement frequency. Their operational principle – measuring distances in a single horizontal or slightly inclined plane – make LiDAR sensors exceptionally stable under varying lighting conditions and suitable for ad-hoc obstacle detection. Standards such as ISO 3691-4 [13], ISO 13855 [14] and VDI 2510 [15] reinforce their relevance, as they are part of mandatory safety regulations for AMR. However, the fundamental limitation of 2D laser scanners arises from their functionality: they cannot sense objects below or above the scanner plane. Near-ground objects mainly fall below the laser's detection height (up to approx. 20 centimetres, depending on the manufacturer's safety-certified installation instructions to avoid collisions regarding the proximity of the lying human body). Even when the scanning plane is angled downward, AMR operate with ground-plane filtering or rely on map-based matching, since the ground-plane would be detected as obstacle provoking a safety stop [8]. As a result, 2D-LiDAR are inappropriate to detect the near-ground objects. This intrinsic sensing gap motivates the integration of 3D sensors.

3D-LiDAR systems provide richer geometric information than their 2D counterparts. High-channel LiDAR systems enable dense 3D reconstruction, supporting obstacle detection, SLAM, and semantic mapping in autonomous vehicles and increasingly in AMR research [16,17]. Their ability to perceive the environment independently of lighting conditions make them promising for challenging industrial settings. Nonetheless,

3D-LiDAR remain challenging for near-ground perception. The key challenge lies in the amount of data and the handling of noise for accurate reconstruction of the 3D environment [18,19]. This is countered by limited computing power and the need for efficient battery use, which needs to ensure the longest possible operational life of AMR. Additionally, even high-channel LiDAR exhibit gaps between laser beams that exceed the height of many near-ground objects. In this case, near-ground obstacles do not intersect the LiDAR beams, and remain undetected [19]. Additionally, LiDAR range measurements near the floor are prone to multipath reflections, occlusion, blurring, noise, and grazing-angle artefacts – effects significantly amplified on polished concrete or metallic surfaces [9]. To overcome these issues, commercial 3D pipelines for AMR apply the ground plane filtering for point clouds from the ground of up to 20 centimetres [20]. While effective for free-space estimation, these methods also eliminate the objects near the ground that obstruct AMR wheels or conveyor mechanisms. While 3D-LiDAR offers substantial advantages over 2D systems and a higher degree of environmental perception, its resolution limits and the required on-board computing resources for data preprocessing prevent reliable detection of near-ground obstacles.

To overcome the constraints of laser-based sensing, additional cameras have become increasingly prevalent. RGB cameras provide high-resolution texture information and are widely used for semantic scene understanding using deep learning approaches. To analyse scenes, modern detectors such as Faster R-CNN [21], YOLO [22] and transformer architectures [23] e.g. DETR [24] and RT-DETRv2 demonstrate improved performance in object detection tasks. However, 2D camera vision suffers from limitations since near-ground objects mostly occupy only few pixels – often below the effective detection threshold of neural networks [8,9]. Stereo and time-of-flight (ToF) cameras aim to provide geometric cues. However, depth accuracy decreases sharply at grazing angles and in lower parts of the field of view. In addition, the problem of precise evaluation of the point cloud and pre-processing of noise for accurate, dimensionally correct isolation of the objects remains challenging – as with 3D LiDAR sensors. However, RGB-D cameras provide a promising extension to LiDAR, yet an efficient near-ground classification pipeline for AMR use remains undeveloped.

Table 1: Taxonomy of the Near-Ground Perception Problem.

Category	Problem Description
Sensor-Level Limitations	Near-ground objects remain below the detection height of 2D LiDAR and fall into beam gaps of 3D LiDAR. Depth sensors exhibit artefacts at grazing angles, including invalid returns, multipath reflections and severe noise. These limitations create structural blind zones on industrial floors, preventing reliable geometric capture of small objects.
Algorithmic Constraints	Detection pipelines lack with small objects, often occupy only a few pixels or sparse point-cloud samples. As a consequence, objects fall below the effective detection threshold of models. Moreover, available training datasets do not reflect industrial near-ground variability, leading to poor generalisation and high uncertainty in floor-level classification.
Environmental Influences	Industrial environments amplify perception degradation through shadows, specular reflections from polished floors, floor-texture variability, temporary clutter, dust, and robot-induced vibrations. These factors distort both RGB and depth data, introduce inconsistent feature representations, and can cause non-deterministic model behaviour during AMR motion.

3. Methodology of the requirement analysis

This research follows a three-stage methodological approach to derive a requirement framework for near-ground object detection and classification for AMR. This procedure produced a coherent, hierarchical structure of requirements. The underlying methodical research process is visualized in Figure 2.

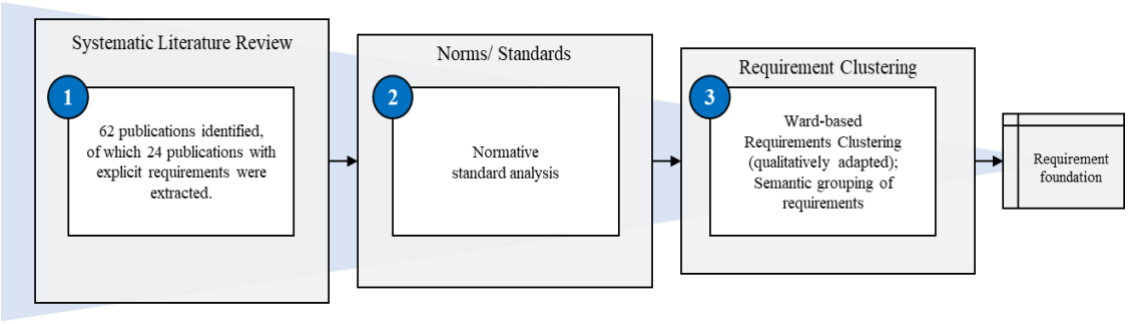


Figure 2: Methodical process of the requirement analysis.

In the first step, a structured literature search identified 62 peer-reviewed publications from 2018 until today, partly published in [5]. Based on the identified publications, 24 papers explicitly formulated technical, environmental, or operational requirements for AMR perception. All requirements were extracted and transferred into an initial requirement set. Secondly, globally relevant safety and robotics standards were analysed to identify safety-related constraints affecting AMR perception. These normative requirements are merged with the literature-derived set to ensure technological completeness. Finally, to organize the combined requirement set, a hierarchical clustering approach based on Ward's method [25] is applied. While Ward's original method minimizes within-cluster variance for numerical data [25], its underlying principle was qualitatively adapted using semantic similarity. Requirements with high conceptual similarity were grouped, and cluster boundaries were defined to maximise thematic distinctiveness.

4. Requirements for near-ground object detection

The following section consolidates the resulting requirements for near-ground object detection and classification for AMR into three requirement clusters: Section 4.1 represents the *operational requirements* implied by the use case of AMR, Section 4.2 shows the *technical requirements* concerning sensing, data processing, and algorithms, and Section 4.3 illustrates the *environmental requirements* reflecting the conditions of industrial shop floors. These clusters are conceptually distinct but tightly interdependent.

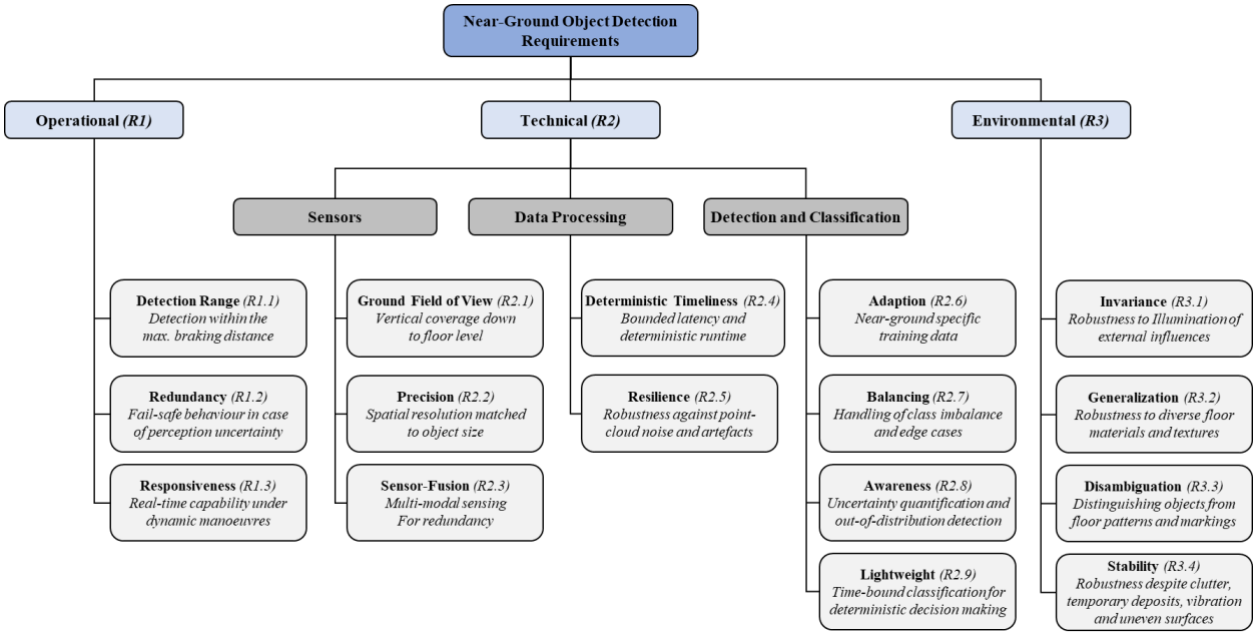


Figure 3: Requirements for near-ground object detection

Operational requirements define what the perception system must achieve in the context of AMR tasks; technical requirements specify how this can be realised in terms of hardware and software; environmental requirements define under which conditions the required performance must be maintained. Figure 3 visualizes the identified requirements for near-ground object detection and classification.

4.1 Operational requirements

Operational requirements specify how the near-ground perception system must support the safe and efficient execution of AMR tasks in factory environments. They directly follow from the mission profile (transport, docking, loading/unloading), the interaction with humans and other equipment. The primary operational requirement is that near-ground objects which can lead to a loss of controllability or collision must be detected with sufficient anticipation to trigger an appropriate reaction of the AMR.

Table 2: Operational Requirements

R1.1	***Detection Range:*** *Detection within the max. braking distance*	The perception system must reliably detect near-ground objects within a distance greater or equal to the dynamic braking distance of the AMR. Formally, the minimal detection range d_{det} must satisfy the following equation: $$d_{det} \geq d_{brake}(v_{AMR}, a_{dec}) + v_{AMR} * t_{lat}$$ where v_{AMR} represents the velocity, a_{dec} is the achievable deceleration, and t_{lat} the perception latency. This requirement is normatively aligned with DIN ISO 13855 [26] coupling machines' motion to response times.
R1.2	***Redundancy:*** *Fail-safe behaviour in case of perception uncertainty*	If the near-ground perception system cannot classify an object with sufficient confidence, the AMR must default to a conservative behaviour such as speed reduction, controlled stop, or re-planning of the path. This requirement reflects the safety principles of the Machinery Directive 2006/42/EC [27] and ISO 3691-4 [28] emphasizing risk reduction even under uncertainty. From a requirements perspective, it implies the definition of confidence thresholds and associated behavioural rules.
R1.3	***Responsiveness:*** *Real-time capability under dynamic manoeuvres*	The perception-decision-actuation loop must be capable of operating in real-time under changing trajectories. When trajectory changes occur late, previously unseen floor regions may become safety-relevant at short notice. Hence, the perception system must maintain a minimal update rate that is consistent with the maximum rotational and translational dynamics of the AMR platform.

4.2 Technical requirements

Technical requirements specify the properties of the sensing hardware, data processing pipeline, and algorithms to fulfil the operational requirements in Section 4.1. They are shaped by the physical limitations of sensors and by the algorithmic constraints of perception methods.

Table 3: Technical Requirements

Sensors	***R2.1***	***Field of View:*** *Vertical coverage down to floor level*	At least one sensor modality must provide reliable measurements down to the floor surface. For LiDAR, this implies installing sensors at heights and angles that ensure intersection with typical near-ground objects. For cameras and depth sensors, this implies a field of view including a sampling of the floor region in front of the AMR.

	R2.2	***Precision:*** *Spatial resolution matched to object size*	The sensing configuration must ensure that smaller near-ground objects occupy a sufficient number of pixels (for cameras) or points (LiDAR or RGB-D cameras) to permit detection and classification. The literature emphasises that near-ground objects often occupy only a few pixels, falling below detection thresholds [8].
	R2.3	***Sensor-Fusion:*** *Multi-modal sensing for redundancy*	Given that each modality has characteristic failure modes (e.g. LiDAR artefacts on reflective floors, camera saturation, invalid depth returns), the requirements analysis supports a multi-modal sensor design. At least two complementary sensing modalities (e.g. LiDAR and RGB-D) should be considered for safety-relevant applications. This requirement is motivated by the diverse failure patterns observed in 2D, 3D, and camera-based systems [29,30].
Data (pre-) processing	***R2.4***	***Deterministic Timeliness:*** *Bounded latency and deterministic runtime*	The perception pipeline must guarantee a worst-case latency t_{lat} from sensor measurement to a stable classification result. As expressed in ***R1.1***, this latency directly enters the safety distance equation. Latency thresholds must be derived from maximal AMR speeds and braking capabilities; this in turn constrains computational complexity and visions' model sizes.
	R2.5	***Resilience:*** *Robustness against point-cloud noise and artefacts*	The near-ground object detection pipeline must incorporate mechanisms for filtering noise. This can be realised via temporal filtering, sensor fusion, or explicit artefact-aware training. The need for such mechanisms follows directly from the degradation phenomena described in Section two (e.g. reflective floors, vibrations, irregular surfaces) [29,31,9].
Algorithms for detection and classification	***R2.6***	***Adaptation:*** *Near-ground specific training data*	Detection and classification models must be trained (or at least fine-tuned) on datasets that contain near-ground objects in realistic industrial contexts, including variations in lighting, floor materials, clutter, and motion artefacts. Generic datasets (e.g. street scenes or household objects) are insufficient, as they do not capture the characteristics of industrial objects.
	R2.7	***Balancing:*** *Handling of class imbalance and edge cases*	Safety-critical objects (e.g. a small metal tool lying on a ramp) may be rare in training data yet must be detected reliably. The perception system must therefore include mechanisms to mitigate class imbalance (e.g. re-weighting, data augmentation) and to handle rare but critical classes.
	R2.8	***Awareness:*** *Uncertainty quantification and out-of-distribution detection*	Given the variability of industrial environments, the perception system must encounter objects and conditions not covered by the training data (e.g. new packaging types, temporary installations, etc..). Requirements must include handling of predictive uncertainty, e.g. via confidence scores or calibrated probabilities.
	R2.9	***Lightweight:*** *Time-bound classification for deterministic decision making*	Lightweight, time-bound classifiers provide faster inference, making them suited for safety-critical perception. By omitting region-proposal stages, they ensure predictable runtimes and reduced computational overhead – key prerequisites for real-time braking-distance reasoning and deployment on embedded AMR hardware.

4.3 Environmental requirements

Environmental requirements describe the range of physical conditions under which near-ground perception must maintain its performance. They correspond to the environmental factors identified in the literature: lighting, shadows, floor reflectivity, particulate contamination, and mechanical vibrations [32,9,33].

Table 4: Environmental Requirements

R3.1	***Invariance:*** *Robustness to Illumination of external influences*	The perception system must remain reliable across varying illumination, including e.g. bright window zones and moving shadows. It must handle specular reflections from polished, metallic, or wet floors introducing artefacts in RGB and depth data [9]. These conditions impose requirements on exposure control, invariance, and robustness to glare – potentially involving optimized camera placement, or fusion with LiDAR.
R3.2	***Generalization:*** *Robustness to diverse floor materials and textures*	The perception system must handle a spectrum of floor materials commonly found in factories (e.g. concrete, epoxy coatings, metal plates, anti-slip mats, tiles) with different textures and colors. Variations in albedo and texture should not cause systematic biases in segmentation or classification. Requirements should therefore include testing and validation across representative floor types.
R3.3	***Disambiguation:*** *Distinguishing objects from floor patterns and markings*	Industrial floors often include painted lines, markers, QR codes, or embedded cable ducts. These patterns may visually resemble small objects at low resolution. The perception system must distinguish between structural floor patterns (which are traversable) and obstacles such as cables. This requirement follows from the observation that many algorithms treat small anomalies on the floor as noise to be filtered [31], while near-ground safety demands their explicit recognition.
R3.4	***Stability:*** *Robustness despite clutter, temporary deposits, vibration and uneven surfaces*	Factories can exhibit varying levels of clutter: temporarily stored packaging, small parts, and other debris along transport routes. Near-ground perception must remain reliable and degrade gracefully under high clutter density. AMR also often traverse floor drains, ramps, and uneven surfaces that induce vibrations, causing distortions and depth artefacts. The system must therefore operate within vibration, supported by e.g. temporal filtering and motion-compensation algorithms.

5. Discussion

The analysis shows that near-ground perception is still a weakness of current AMR systems. Although AMR employ 2D and 3D LiDAR and RGB-D cameras, their structural limitations – blind spots near the floor, sub-pixel object size, depth artefacts, glare, shadows and vibration – prevent reliable detection of small objects commonly found in industrial environments [5,8,9].

The three requirement clusters reveal a clear mismatch: safety standards demand detection within braking distance and full footprint coverage, while commercial perception stacks intentionally filter out the floor region. Technical constraints (e.g. sensor geometry, latency, training data, etc…) and environmental challenges (e.g. reflective floors, clutter, dust, etc…) further reinforce why such obstacles remain difficult to detect. Furthermore, the operational, technical and environmental requirement clusters are conceptually separated but strongly coupled in practice. Operational safety requirements in Section 4.1 translate into constraints on detection range, latency, and classification reliability, which must be satisfied by the technical design, presented in Section 4.2, under the environmental conditions characterised in Section 4.3.

Overall, the requirement analysis illustrates that near-ground perception cannot be treated as a minor extension of standard obstacle detection but requires dedicated sensing, tailored algorithms and data processing pipelines and explicit robustness against industrial conditions.

6. Conclusion and future research activities

This paper provides a requirement framework for near-ground object detection in AMR, integrating operational, technical, and environmental constraints. The findings from the analysis confirm that existing AMR sensor configurations and generic deep-learning models are not sufficient for robust floor-level perception in industrial settings. One of the main challenges in this regard is the handling and preprocessing of large amounts of data in real time while ensuring sufficient accuracy for filtering noise and obtaining the closest possible 3D representations of the near-ground objects. Furthermore, a subsequent challenge following object recognition is the integration into the control level of the AMR as well as resource-efficient data management to enable AMR to remain in operational mode as long as possible. In this regard, the authors are conducting further research to bridge this research gap and enable improved utilization of AMR systems offering a more advanced level of perception.

References

[1] Fottner, J., Clauer, D., Hormes, F., Freitag, M., Beinke, T., Overmeyer, L., Gottwald, S.N., Elbert, R., Sarnow, T., Schmidt, T., Reith, K.B., Zadek, H., Thomas, F., 2021. Autonomous Systems in Intralogistics – State of the Art and Future Research Challenges. Logistics Research, 1–41.

[2] Fragapane, G., Koster, R. de, Sgarbossa, F., Strandhagen, J.O., 2021. Planning and control of autonomous mobile robots for intralogistics: Literature review and research agenda. European Journal of Operational Research 294 (2), 405–426.

[3] Niloy, M.A.K., Shama, A., Chakrabortty, R.K., Ryan, M.J., Badal, F.R., Tasneem, Z., Ahamed, M.H., Moyeen, S.I., Das, S.K., Ali, M.F., Islam, M.R., Saha, D.K., 2021. Critical Design and Control Issues of Indoor Autonomous Mobile Robots: A Review. IEEE Access 9, 35338–35370.

[4] Hofmann, C., Taliercio, F., Walter, J., Franke, J., Reitelshöfer, S., 2023. Towards Adaptive Environment Perception and Understanding for Autonomous Mobile Robots. 2023 IEEE Symposium Sensor Data Fusion and International Conference on Multisensor Fusion and Integration (SDF-MFI), 1–8.

[5] Lackner, T., Hermann, J., Kuhn, C., Palm, D., 2024. Review of autonomous mobile robots in intralogistics: state-of-the-art, limitations and research gaps. Procedia CIRP 130, 930–935.

[6] Badrloo, S., Varshosaz, M., Pirasteh, S., Li, J., 2022. Image-Based Obstacle Detection Methods for the Safe Navigation of Unmanned Vehicles: A Review. Remote Sensing 14 (15), 1–26.

[7] Nguyen, N.-D., Do, T., Ngo, T.D., Le, D.-D., 2020. An Evaluation of Deep Learning Methods for Small Object Detection. Journal of Electrical and Computer Engineering 2020, 1–18.

[8] Takeda, K., Tanaka, K., Nakamura, Y., 2022. Small Object Change Detection for Small Obstacle Avoidance in Everyday Robot Navigation. Research paper, Tokio, 5 pp.

[9] Wei, Z., Zou, Y., Xu, H., Wang, S., 2025. Small Object Detection in Traffic Scenes for Mobile Robots: Challenges, Strategies, and Future Directions. Electronics 14 (13), 1–21.

[10] Yu, C., Rastogi, C., Zhou, Z., Norris, W.R., 2023. A Self-Supervised Miniature One-Shot Texture Segmentation (MOSTS) Model for Real-Time Robot Navigation and Embedded Applications.

[11] Hakanen, T., Kemppi, P., Tikka, P., 2023. Mobile 3D LiDAR-based Object and Change Detection in Production and Operations Management. 19th International Conference on Autonomic and Autonomous Systems, ICAS 2023, 1–7.

[12] Zhu, A.Z., Thakur, D., Ozaslan, T., Pfrommer, B., Kumar, V., Daniilidis, K., 2018. The Multi Vehicle Stereo Event Camera Dataset: An Event Camera Dataset for 3D Perception. IEEE Robot. Autom. Lett. 3 (3), 2032–2039.

[13] Deutsches Institut für Normung e. V. (DIN). Flurförderzeuge – Sicherheitstechnische Anforderungen und Verifizierung – Teil 4: Fahrerlose Flurförderzeuge und ihre Systeme. Beuth-Verlag, Berlin.

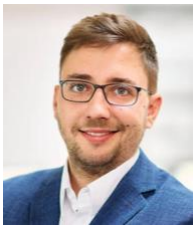
[14] Deutsches Institut für Normung e. V. (DIN). Sicherheit von Maschinen – Anordnung von Schutzeinrichtungen im Hinblick auf Annäherung des menschlichen Körpers. Beuth-Verlag, Berlin, 203 pp.

[15] Verein Deutscher Ingenieure e.V. (VDI), 2022. VDI 2510 Fahrerlose Transportsysteme (FTS) - Sicherheit von FTS. Beuth-Verlag, Berlin.

[16] Jens Behley, Martin Garbade, Andres Milioto, Jan Quenzel, Sven Behnke, Cyrill Stachniss, Jurgen Gall, 2019. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 9297–9307.
[17] Wozniak, M.K., Kårefjärd, V., Hansson, M., Thiel, M., Jensfelt, P., 2023. Applying 3D Object Detection from Self-Driving Cars to Mobile Robots: A Survey and Experiments. 2023 IEEE International Conference on Autonomous Robot Systems and Competitions, 3–9.
[18] Balta, H., Velagic, J., Bosschaerts, W., Cubber, G. de, Siciliano, B., 2018. Fast Statistical Outlier Removal Based Method for Large 3D Point Clouds of Outdoor Environments. IFAC-PapersOnLine 51 (22), 348–353.
[19] Zermas, D., Izzat, I., Papanikolopoulos, N., 2017. Fast segmentation of 3D point clouds: A paradigm on LiDAR data for autonomous vehicle applications. 2017 IEEE International Conference on Robotics and Automation (ICRA), 5067–5073.
[20] Badrloo, S., Varshosaz, M., Pirasteh, S., Li, J., 2022. Image-Based Obstacle Detection Methods for the Safe Navigation of Unmanned Vehicles: A Review. Remote Sensing 14 (15), 1–26.
[21] Ren, S., He, K., Girshick, R., Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.
[22] Redmon, J., Farhadi, A. YOLOv3: An Incremental Improvement.
[23] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2020. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, 22 pp. https://arxiv.org/pdf/2010.11929v2.
[24] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S., 2020. End-to-End Object Detection with Transformers.
[25] Ward Jr., J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58 (301), 236–244.
[26] Deutsches Institut für Normung e. V. (DIN). Sicherheit von Maschinen – Anordnung von Schutzeinrichtungen im Hinblick auf Annäherung des menschlichen Körpers. Beuth-Verlag, Berlin, 203 pp.
[27] Europäische Union, 2006. Richtlinie 2006/42/EG des Europäischen Parlaments und des Rates über Maschinen und zur Änderung der Richtlinie 95/16/EG (Neufassung). Europäischen Parlaments und des Rates, Brüssel. Accessed 27 October 2024.
[28] Deutsches Institut für Normung e. V. (DIN). Flurförderzeuge – Sicherheitstechnische Anforderungen und Verifizierung – Teil 4: Fahrerlose Flurförderzeuge und ihre Systeme. Beuth-Verlag, Berlin.
[29] Eppenberger, T., Cesari, G., Dymczyk, M., Siegwart, R., Dube, R., 2020. Leveraging Stereo-Camera Data for Real-Time Dynamic Obstacle Detection and Tracking. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 10528–10535.
[30] Xu, Z., Zhan, X., Xiu, Y., Suzuki, C., Shimada, K., 2024. Onboard Dynamic-Object Detection and Tracking for Autonomous Robot Navigation With RGB-D Camera. IEEE Robot. Autom. Lett. 9 (1), 651–658.
[31] Steinke, N., Goehring, D., Rojas, R., 2024. GroundGrid: LiDAR Point Cloud Ground Segmentation and Terrain Estimation. IEEE Robot. Autom. Lett. 9 (1), 420–426.
[32] Brand, L., Wunderle, Y., Hohmann, S., 2024. Ultrasonic Object Detection and Classification for AMR Safety. 2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation, 1–6.
[33] Wunderle, Y., Lyczkowski, E., Hohmann, S., 2024. Safe object detection in AMRs - a Survey. IEEE 33rd International Symposium on Industrial Electronics (ISIE), 1–8.

Biography

Thorge Lackner works as a research associate at Reutlingen University's Faculty of Sustainability and Technology (NXT). He holds a Master of Science in Industrial Engineering from the ESB Business School and a Master of Engineering in Engineering Management from Stellenbosch University in South Africa. He is currently pursuing his doctoral degree at Reutlingen University.



Dr. Julian Hermann works as a senior researcher at Reutlingen University's Faculty of Sustainability and Technology (NXT). After completing his studies in Mechatronics and Industrial Engineering at Reutlingen University, he earned his PhD from Stellenbosch University in South Africa. His research focuses on AI in production, mobile robotics and connectivity technologies such as 5G.

Christian Kuhn works as a research associate at Reutlingen University's Faculty of Sustainability and Technology (NXT). He holds a Master of Science in Industrial Engineering from the ESB Business School. His work focuses on AI-driven production solutions, digital twins for machines, and teaching and learning concepts within the area of digital transformation.

Prof. Dr.techn. Daniel Palm is a Professor of Logistics Management at Reutlingen University's Faculty of Sustainability and Technology (NXT). He heads the MSc "Digital Industrial Management and Engineering", directs the Reutlingen Centre Industry 4.0 as well as Werk150, university's learning factory and heads the Fraunhofer Institute of Manufacturing Engineering and Automation branch in Reutlingen. His work focuses on AI-driven production and logistics, digital supply chains, and Industry 4.0 transformation.

8th Conference on Production Systems and Logistics

Selection Of Indicators For Integrating Ecological Sustainability Into Product Portfolio Management

Nikolai Kelbel[1], Alexander Keuper[1], Günther Schuh[1, 2]

[1]*Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, 52062 Aachen, Germany*
[2]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

Sustainability has become an increasingly important factor in decision-making of companies over the past decades. To comply with regulations, companies need a systematic controlling approach for identifying weak points, creating transparency, and supporting the definition of improvement measures for their product portfolio. Selecting an appropriate set of controlling indicators is crucial to ensure a successful transformation towards increased sustainability throughout a company's product portfolio. While existing approaches examine the sustainability indicator selection from various specific perspectives, they lack both holism and the consideration of requirements specific to product portfolio management (PPM), which decreases their practical relevance in this context. Therefore, this paper aims to systematically select a sustainability indicator set for PPM by conducting a meta study of existing literature. This is done by extracting relevant approaches from a systematic literature review and consolidating their proposed indicator sets and perspectives. The set is gradually narrowed down by evaluating the formal and content wise quality of the indicators as well as their usability in practice. The proposed methodology is shown to streamline the definition of sustainability controlling in practice.

Keywords

Product portfolio management; Sustainability; Controlling; Indicators; KPI

1. Introduction

Sustainability has evolved into a central concern for companies across industries, driven by increasing regulatory requirements and heightened expectations from diverse stakeholder groups [1]. While sustainability is often discussed as a holistic concept encompassing economic, social, environmental, and governance dimensions, the ecological aspect in particular presents significant complexity [2]. Ecological sustainability cannot be reduced to greenhouse gas emissions alone; it also comprises factors such as water consumption, biodiversity impacts, land use, and toxicity [3]. This breadth of relevant dimensions results in a challenging landscape for companies seeking to operationalize sustainability within their management processes.

A clear definition of relevant ecological dimensions is essential for integrating sustainability into established business practices such as product portfolio management (PPM). However, existing research offers limited guidance in this context [6,4,5]. Most frameworks propose sets of ecological indicators tailored to product development or production but seldom address the specific requirements of PPM. Thus, there is a pressing need for a robust indicator set that enables manufacturing companies to systematically steer the sustainability

DOI: https://doi.org/10.15488/21014
ISSN: 2701-6277

of their product portfolios. This set should support compliance with regulatory demands and address stakeholder expectations while maintaining profitability and operational viability.

This paper seeks to address this by systematically selecting a sustainability indicator set for PPM by conducting a meta study of existing literature. In doing so, it aims to facilitate the effective embedding of ecological sustainability into strategic decision-making. In chapter 2 the fundamentals and related research are elaborated. Chapter 3 presents the proposed methodology, chapter 4 presents the resulting indicator set before chapter 5 sums up the paper's results and gives an outlook.

2. Fundamentals and Related Work

In the following the fundamentals of sustainability-oriented controlling are presented. Subsequently, related work is presented.

2.1 Sustainability-oriented Controlling

Originally, controlling is a term from internal accounting and includes the functions of informing, steering, controlling and planning [7,8]. The central aspect of it is to provide a comparison between the status quo and the targets [9]. Defining relevant indicators is a proven method for measuring target achievement. These indicators are quantitatively measurable variables that are defined individually for each target and can be influenced by the company [10]. When indicators that are interrelated are grouped together, this is referred to as an indicator system [8], or "set" in the context of this paper. This enables to measure multiple target dimensions.

Due to the sustainability transition in the industry the classic controlling is expanded by sustainability aspects [11]. As part of its information function, controlling must ensure that relevant information is available in sufficient quantity, with the necessary accuracy and certainty. In addition to analyzing opportunities and risks, sustainability-oriented controlling (SOC) also involves developing potential strategies for ensuring target achievement [11]. Sustainability targets can for example originate from regulatory or customer requirements and involve material reductions, fostering renewable energies or the phase out of environmentally detrimental products [12,13]. The main difficulty lies in the multidimensional nature of the decision-making criteria. While financial criteria are generally monetary in nature, material and energy flows are represented in units of quantity on the ecological side [11]. SOC has various instruments at its disposal for evaluating and controlling which explicitly take sustainability aspects into account. One of the most established instruments is the Life Cycle Assessment (LCA) [14].

Due to its interface function, its linkage to strategic management and its responsibility for a company's products, SOC falls into the domain of PPM [15,16]. PPM is a component of strategic management. It ensures the long-term profitability of a company by aligning its product range with strategic goals and objectives [16]. This involves strategic choices, resource allocation, project selection, and balancing to ensure the portfolio success [17].

2.2 Related Work

Various authors have already worked on indicator selection in the sustainability context. The following gives a short overview of significant studies.

In their study, ISSA ET AL. [18] identify the challenges companies face when selecting and applying environmental performance indicators and emphasize the lack of a structured approach to their selection. Their five-step process includes the following steps: setting environmental priorities and goals, preselecting indicators based on defined criteria, selecting indicators from the database, adapting and creating new indicators, and implementing the selected indicators. The study is limited by its focus on product development. In addition, the selection is rather qualitative in nature, only suggesting sample questions for

the final selection on a company-specific basis. Finally, the implementation process is only described superficially.

PARK AND KREMER [19] address the problem of insufficient transparency by presenting a data-driven approach using text mining to categorize sustainability indicators. Through a survey of companies in the manufacturing and service sectors, the authors determine the industry's perception of the practicality and usefulness of the identified indicators. Ultimately, the authors identify a set of 18 indicators that meet their requirements for practicality and usefulness. However, the selection of indicators is based exclusively on the literature available up to 2016, excluding more recent developments. In addition, the assessment of the practicality and usefulness of the indicators is based on qualitative user surveys, which, however, require a certain level of maturity in the use of such indicators. If users are lacking experience in the field of sustainability controlling, the results of such a survey must be questioned.

Mesa et al. [20] emphasize the need for specific sustainability indicators for product families in the context of the circular economy. The authors criticize that existing indicators are mainly focused on individual products and do not sufficiently consider the product family approach, in which common product platforms and modules generate different product variants. To close this gap, the authors develop a set of six indicators that focus on material flow, reusability, reconfiguration, and functional performance. The approach refers to the level of product families and is therefore potentially transferable to portfolios. However, the focus is only on indicators related to the circular economy. In addition, the indicators are based on idealized assumptions about the potential for reuse and recycling, without taking into account the actual operationalization of such measures.

SHERIF ET AL. [5] address the challenge manufacturing companies face in selecting and interpreting suitable sustainability indicators to transform their processes. Their work includes the development of a database for sustainability indicators, their categorization, evaluation, and the compilation of a final set of indicators for specific goals. The proposed categorization is based on sustainability goals and manufacturing operations, thereby achieving a comprehensive mapping of indicators in relation to production processes but neglecting the sustainability of products or product portfolios.

The diversity of the approaches presented and their temporal dispersion demonstrate the continued relevance of the topic. However, this also shows that research has not yet been able to define a generic, practical set of indicators. Furthermore, most approaches relate either to product development or to production. No approach could be identified that relates to the management of product portfolios. All approaches take a similar approach to the selection of indicators. They are first compiled based on literature references or expert interviews and then examined and shortened in terms of their usefulness and practicability. More detailed instructions on the specific selection of indicators, as outlined qualitatively by ISSA ET AL. [18], are also lacking in the literature. Without relevant expert knowledge, it is therefore not possible to transfer research results into practice to a sufficient degree. Thus, there is still a need for research into a structured and detailed selection of indicators relevant to PPM in the context of sustainability. The research question is formulated as follows: *How can the ecological sustainability of products be measured in product portfolio management?*

3. Methodology

There are already numerous approaches in the existing literature that have compiled indicator sets from different perspectives using different methods. However, as explained in the previous chapter, the PPM perspective is lacking. This paper closes this gap by building on existing approaches and following VOM BROCKE's idea of "standing on the shoulders of giants" [21] through a meta-study.

This is done by analyzing existing literature using structured literature research. For this purpose, the Scopus database is searched for relevant literature using the following search string.

TITLE-ABS-KEY ((sustainability OR sustainable OR ecological OR ecologic OR environment OR environmental OR (circular economy)) AND (KPI OR indicator OR criteria) AND (product OR manufacturing) AND (categorization OR categorizing OR set OR selection OR guide OR database OR identification))

The results are sorted by relevance and the first 200 results are analyzed. Furthermore, the results are sorted by novelty and the first 100 results are added to the analysis. The relevance of the results is checked by reading the elements of the articles in the following order: title, abstract, conclusion, full text. The quality and relevance of the articles is checked using the PRISMA checklist [22]. The indicator sets of the resulting articles are then transferred to a long list, duplicates are removed, and the list is consolidated. The main part of the consolidation process is the evaluation of the formal quality of the indicators. The evaluation criteria are based on the existing indicator evaluations of SHERIF ET AL. [5] and ISO 22400 [23]. The result is a qualitative evaluation that is operationalized via descriptions. The procedure is visualized in Figure 1.

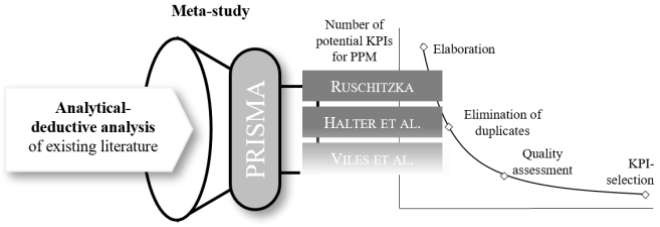


Figure 1: Research approach of the meta-study

4. Results

The present study deliberately draws on earlier contributions in the field of sustainability indicator selection. One approach, in particular, distinguishes itself through both its holistic orientation and its methodological rigor, RUSCHITZKA [24]. Her research develops a function-oriented method for analyzing the ecological performance of products in mechanical and plant engineering in order to identify levers for improving product sustainability. As a part of this, she reviews sustainability-related indicators and documents their frequency of occurrence in academic literature. In the process, she focuses primarily on function-oriented ecological analysis at the inventory level, because the use of midpoint and endpoint indicators - common in many LCAs - can obscure direct causal relationships and, hence, the identification of concrete improvement measures. To avoid this limitation and to ensure greater transparency for product development, she carries out a systematic literature review that compiles 1 064 ecological indicators from 101 sources. Following the consolidation logic proposed by WELSING [25], synonymous or only slightly divergent designations are merged, yielding 246 ‘generic’ ecological indicators. These are subsequently classified into the categories “energy”, “emissions”, “materials”, “waste”, “water and wastewater”, and “other”. Indicators related to land use and biodiversity are omitted owing to their lack of direct product relevance. A frequency-based selection results in a shortlist of 30 indicators, which together account for approximately 60 % of all mentions. RUSCHITZKA [24] explicitly notes that the compiled lists are not exhaustive and may be expanded over time. Moreover, she confines the scope to product-specific environmental impacts and the corresponding life-cycle processes. In this way, she provides a robust basis for assessing ecological aspects in product development while keeping the framework open to future extensions.

Despite its breadth, RUSCHITZKA’s indicator set is only partially suitable for application in PPM. Excluding midpoint indicators is appropriate when direct improvement measures are to be derived at the product-development level; at the portfolio level, however, a strategic, aggregated perspective is equally indispensable. Consequently, her indicator set is adopted here merely as a foundation and is systematically extended by integrating multiple perspectives as a result of the present meta-study. They are elaborated in the following.

The ELLEN MACARTHUR FOUNDATION [26] developed an approach to measure the circularity of products by defining various indicators. The indicators focus on material flow and product circulation and aggregate multiple levels of sustainability. Because of its novel perspectives and aggregated indicators, the contribution is considered in this meta-study.

HALTER ET AL. [27] shift the focus to circular economy indicators. Employing a structured PRISMA-based literature review, they identify 798 indicators across 695 publications, having first excluded aggregated indicators. The indicators are categorized by hierarchy level, sustainability pillar and product-life-cycle stage. Logical consolidation of semantically similar indicators yields 67 generalized indicators. In addition, the authors propose an extended SMART method for the qualitative evaluation of indicators and as a decision-support tool when compiling bespoke indicator sets. Because of their exclusive focus on circular economy indicators and their temporal origin, the findings of HALTER ET AL. are included in the present meta-study.

Viles et al. [28] contribute yet another perspective. On the basis of a structured literature review, they compile an indicator list that is derived from, and subsequently categorized according to, the United Nations Sustainable Development Goals (SDGs). The assignment of indicators to SDGs is validated by means of semi-structured expert interviews. After removing indicators that are not environmentally relevant, a list of 39 indicators remains; this list is incorporated into the present meta-study owing to its novel perspective.

Although already contained in Ruschitzka's compilation, the approach of Park and Kremer [19] broadens the perspective of the meta-study and is therefore considered independently of her quantitative analysis. The authors employ a text-mining approach to identify indicators and further screen them for usefulness and practicability by means of logistic regression. A more detailed account of their method is provided in Chapter 2.2.

ALJAMAL ET AL. [6] add a further layer by analyzing ecological indicators via social-network analysis. The outcome is a set of quantitative and qualitative indicators that occupy central and influential positions within the indicator network and are therefore of high relevance for measuring ecological sustainability. Owing to the recency of its publication, this work has not yet been incorporated into any aggregated study.

Finally, GAUTAM ET AL. [29] are also considered. Their approach resembles that of RUSCHITZKA. However, focusing on India's manufacturing industry, the authors introduce an international dimension to the analysis. Only the environmentally relevant indicators are retained for further consideration.

The indicator sets yielded by these approaches are consolidated into a longlist and complemented by the midpoint indicators of a LCA [14], as a standardized method for determining product sustainability. The longlist is successively cleared of duplicates and clustered in line with RUSCHITZKA's categories. For assessing the quality of the indicators in the context of sustainability-oriented PPM a qualitative assessment is performed. The corresponding assessment criteria are inspired by SHERIF ET AL. [5], visualized in Figure 2 and described below.

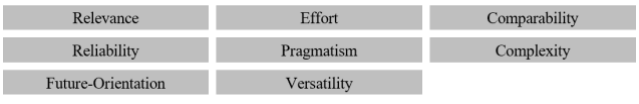


Figure 2: Criteria for assessing the indicator quality

To evaluate the indicators, the criteria are divided into two categories: knockout criteria and evaluative criteria. Knockout criteria are operationalized using a binary, qualitative scale and, if evaluated negatively, result in the exclusion of the indicator. They are listed below.

Relevance describes whether the indicators are relevant for managing a product portfolio in the context of environmental sustainability. Compliance with this criterion is already ensured by the target orientation of the meta study. **Reliability** ensures that the indicator accurately reflects reality, is sufficiently defined, and

does not neglect any important aspects. A rating of "0" means that the indicator is imprecisely defined or does not take important aspects into account.

Evaluative criteria support the comparison of indicator quality and are measured on a qualitative scale from 0 (low quality) to 2 (high quality). They are listed below.

Future-Orientation describes whether an indicator can reflect reality over time and can also be projected into the future. Indicators that can only be collected after a significant delay and cannot be projected into the future are assigned a "0." Indicators with a "2" can be collected in real time and are easy to project. **Effort** describes the financial, time, and personnel costs of continuous indicator collection and updating. A rating of "0" means that the effort barely manageable, while "2" describes indicators that are very easy to collect or are already available. **Pragmatism** checks whether indicators can be linked directly and intuitively to specific improvement measures. A "0" indicates that an indicator can only be influenced with complex packages of measures, whereas an indicator rated "2" can be influenced directly and intuitively with measures. **Versatility** describes the extent to which an indicator can be used to measure multiple objectives or has benefits beyond controlling (e.g. for reporting). A "0" indicates indicators that only have value in a controlling context. Indicators with a "2" enable synergies outside of controlling, for example in reporting. **Comparability** describes the extent to which an indicator can be used to meaningfully compare the performance of multiple products with each other. For this purpose, it is important to consider whether an indicator is relative or absolute in nature. A "0" describes indicators that do not allow for meaningful comparison, while a "2" describes those that do. **Complexity** checks whether an indicator can be easily understood, uniformly measured, and clearly interpreted by stakeholders. A "0" indicates an indicator that is impossible to use without expert knowledge, while a "2" indicates one that is intuitive and directly usable.

After the indicator longlist was reduced through the quality assessment (see Figure 3), seven indicators remain, which are shown in Figure 4 and described below. The selection was made as follows:

- The indicator has a high-quality rating.
- The indicator is relevant for the entire product life cycle.
- Aggregated indicators are preferred over specific ones for the compilation of a generic set.

		Relevance	Reliability	Future-Orientation	Effort	Pragmatism	Versatility	Comparability	Complexity	**Sum**
Energy	…									
	Energy consumption	1	1	2	2	2	2	1	2	11
	Energy efficiency	1	1	2	1	1	2	1	1	8
Emis-sions	CO_2 emissions	1	1	1	0	1	2	1	1	6
	Prevented emissions	1	0	-	-	-	-	-	-	-
	…									

Figure 3: Excerpt from the indicator quality assessment

Energy consumption	Material consumption	Waste generation
THG-Emissions	Water consumption	
Linear Flow Index	Product utility	

Figure 4: Generic indicator-set for the sustainability-oriented PPM

Energy consumption denotes the total amount of energy required by a product throughout its life cycle, from raw-material extraction and manufacturing to usage and end-of-life. For technical products that draw energy during the use phase, especially those with long service lives, this stage can be expected to dominate overall energy demand [30]. Since the user-specific energy mix is rarely known [30], the national average electricity mix must generally be assumed. Under this premise, the indicator may be restricted to the use phase, and the differentiation by energy source can be neglected. In the context of PPM, the indicator controls

a main source of a product's ecological impact and can be connected to various improvement measures across R&D, service or sales. It should, however, be acknowledged that certain products do not fit this pattern. A cast component, for instance, may have no energy demand in use yet requires substantial energy during production; in such cases the indicator must be interpreted differently. Where uncertainty prevails, pilot LCAs on representative products can offer guidance. For transforming the indicator to a relative one, enabling product comparison, it needs to be relativized, e.g. by using the revenue of the product.

Greenhouse gas (GHG) emissions, employed here synonymously with global warming potential, aggregate all climate-relevant emissions and are usually expressed in carbon-dioxide equivalents. Although this indicator is frequently cited in the literature [24], it achieves only moderate quality ratings, primarily because of the considerable measurement effort and limited data availability. Determining life-cycle-wide GHG emissions typically necessitates comprehensive analyses such as full LCAs. Nevertheless, as most alternative indicators in the same category show the same issue, GHG emissions are retained as the most holistic metric. For PPM the indicator plays a vital role as an overarching and easy-to-understand indicator for stakeholders and in the reporting context. Companies that handle hazardous or toxic substances can measure the corresponding emissions separately, for example by monitoring acidification, eco-toxicity, ionizing-radiation impacts or other toxic releases.

Within the material category, two indicators are selected. The **Linear Flow Index** quantifies the proportion of material flowing in a linear fashion which covers recycled material as well as cycled material in the context of circular economy. This indicator is rated highly in terms of quality; its chief limitation lies in incomplete data transparency along the supply chain. It is particularly informative for materials that combine high environmental burdens with favorable recycling or reuse characteristics, such as metals, and can be readily linked to managerial measures. In PPM the indicator's benefit lies in its coverage of product improvements in the context of circular economy. **Material consumption** records the absolute mass of material required for a product. It may include inputs that are consumed during manufacturing but do not remain in the product (e.g. production scrap), as well as materials expended during use (e.g. 3-D-printer filament). The latter component is highly scenario-dependent and therefore subject to considerable uncertainty. Even so, the indicator aligns well with measures aimed at enhancing product efficiency and supplements the previous indicator by providing an absolute value. It can be managed as an aggregated total or broken down by material type or life-cycle stage. Hazardous substances can, in addition, be tracked via dedicated indicators such as chemical consumption, human-toxicity potential or the share of dangerous materials. The circular economy literature proposes numerous further indicators, e.g. circular material use rate or the share of products and packaging designed for circularity, but these indicators are treated here as secondary. Their ultimate purpose is to mitigate environmental impacts, which are already covered by the primary indicators and thus not pursued further.

Water consumption captures the withdrawal of fresh water over the entire product life cycle. The indicator is of good intrinsic quality but entails substantial data-collection effort owing to supply-chain dependencies and varying usage scenarios (e.g. for washing machines). For PPM its relevance depends on whether water is material for the company. Where toxic substances are involved, supplementary control of eutrophication may be advisable.

Waste generation encompasses all material streams that exit the system boundary as waste during the product's life cycle (e.g. non-reused production scrap). For PPM this connects to measures aiming at product and production efficiency. If toxic waste arises, it may be monitored separately.

Finally, the category "other" includes the indicator **product utility**, defined as the product of the length of the product's lifetime and its intensity of use. Although highly rated in the indicator assessment, it likewise entails high data-collection effort. The indicator offers substantial leverage over product design (e.g. maintainability, long-term efficiency) and, when potential second or subsequent life cycles are considered, can drive the implementation of circular economy strategies or business models.

5. Conclusion

For enabling the sustainability transition of companies' product portfolio management must implement the new dimension of (ecological) sustainability within their controlling processes. This paper aims to offer a generic indicator set to foster this without the need for expert knowledge. The presented approach builds on existing research and performs a meta-study as well as an indicator assessment. This results in a set of seven indicators to measure the ecological performance of a product portfolio. All indicators can be connected to steering activities and improvement measures, thus, support PPM in its decision-making. In practice, companies can build their sustainability controlling on this basis. It enables companies, especially those at the beginning of their sustainability transition, to save resources by doing so. The study is limited by its focus on academic literature and by its lack of validation. The authors continue to work on the validation of the set via expert interviews and the incorporation of a practical view via a top-down indicator derivation from actual company targets. Furthermore, future work will be focusing on adapting product portfolio management processes to implement the indicator set into decision-making.

References

[1] IPCC, 2023. Climate Change 2023: Synthesis Report. Summary for Policymakers. IPCC. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_SPM.pdf. Accessed 2 February 2025.

[2] Vandaele, N., Decouttere, C., 2011. Sustainable R&D Portfolio Assessment. SSRN Journal.

[3] Ciroth, A., Arvidsson, R., 2021. Life cycle inventory analysis: Methods and data. Springer, Cham, 209 pp.

[4] Hallstedt, S.I., 2017. Sustainability criteria and sustainability compliance index for decision support in product development. Journal of Cleaner Production 140, 251–266.

[5] Sherif, Z., Sarfraz, S., Jolly, M., Salonitis, K., 2022. Identification of the Right Environmental KPIs for Manufacturing Operations: Towards a Continuous Sustainability Framework. Materials (Basel, Switzerland) 15 (21).

[6] Aljamal, D., Salem, A., Khanna, N., Hegab, H., 2024. Towards sustainable manufacturing: A comprehensive analysis of circular economy key performance indicators in the manufacturing industry. Sustainable Materials and Technologies 40, e00953.

[7] Behringer, S., 2021. Controlling, 2., überarbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden, 165 pp.

[8] Buchholz, L., 2019. Strategisches Controlling: Grundlagen - Instrumente - Konzepte, 3., vollständig aktualisierte Auflage ed. Springer Gabler, Wiesbaden, Heidelberg, 255 pp.

[9] Kümper, T., 2024. Controlling: Eine Einführung mit Storytelling-Ansatz. Springer Gabler, Wiesbaden, Heidelberg, 264 pp.

[10] Weber, W., Kabst, R., Baum, M., 2018. Einführung in die Betriebswirtschaftslehre. Springer Fachmedien Wiesbaden, Wiesbaden.

[11] Wellbrock, W., Ludin, D., Krauter, S., 2020. Nachhaltigkeitscontrolling: Instrumente und Kennzahlen für die strategische und operative Unternehmensführung. Springer Gabler, Wiesbaden, Heidelberg, 52 pp.

[12] European Commission, 2023. Regulation (EU) 2023/2772. European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202302772. Accessed 11/13/2024.

[13] Schuh, G., Keuper, A., Ruschitzka, C., Schümmelfeder, S., Prumbohm, F., 2024. Managing Sustainable Innovations : Ergebnisbericht der Benchmarking Studie. Werkzeugmaschinenlabor WZL der RWTH Aachen, Lehrstuhl für Produktionssystematik. https://publications.rwth-aachen.de/record/984927/files/984927.pdf. Accessed 11/19/2024.

[14] ISO. Environmental management — Life cycle assessment — Principles and framework.

[15] Mohammad, J., Pan, Y.-C., 2022. Sustainability, the fourth pillar of project portfolio management – a holistic approach. Journal of Modern Project Management 9 (2), 198–215.

[16] Riesener, M., Dölle, C., Schuh, G., Lauf, H., Jank, M.-H., 2019. Performance-driven and company goal-orientated design of product portfolios: A methodological framework. Procedia CIRP 84, 725–730.

[17] Cooper, R., 1999. New product portfolio management: practices and performance. J of Product Innov Manag 16 (4), 333–351.

[18] Issa, I.I., Pigosso, D.C., McAloone, T.C., Rozenfeld, H., 2015. Leading product-related environmental performance indicators: a selection guide and database. Journal of Cleaner Production 108, 321–330.

[19] Park, K., Kremer, G.E., 2017. Text mining-based categorization and user perspective analysis of environmental sustainability indicators for manufacturing and service systems. Ecological Indicators 72, 803–820.

[20] Mesa, J., Esparragoza, I., Maury, H., 2018. Developing a set of sustainability indicators for product families based on the circular economy model. Journal of Cleaner Production 196, 1429–1442.

[21] Jan vom Brocke, Alexander Simons, Kai Riemer, Bjoern Niehaves, Ralf Plattfaut, and Anne Cleven, 2015. Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. Communications of the Association for Information Systems 2015 (37).

[22] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Systematic reviews 10 (1), 89.

[23] ISO, 2014. Automation systems and integration — Key performance indicators (KPIs) for manufacturing operations management, 1st ed.

[24] Ruschitzka, C., 2025. Funktionsorientierte Ökologieanalyse im Maschinen- und Anlagenbau, 1st ed. RWTH Aachen University.

[25] Welsing, M., 2023. Bewertung der ökologischen Nachhaltigkeit von Produktionsnetzwerkkonfigurationen. Dissertation, 1. Auflage ed.

[26] Ellen MacArthur Foundation, 2019. An approach to measuring circularity. https://content.ellenmacarthurfoundation.org/m/77e62bc9924c20d0/original/Circularity-Indicators-Methodology.pdf. Accessed 4 January 2026.

[27] Halter, F., Wietschel, L., Thorenz, A., Tuma, A., 2025. Techno-sustainable analysis of circular economy-indicators for corporate supply chains. Journal of Cleaner Production 503, 145078.

[28] Viles, E., Cavallieri, M.S., Kalemkerian, F., Montoya-Torres, J.R., Santos, J., 2024. Assessing sustainable production: A selection of indicators aligned with the SDGs. Sustainable Development 32 (3), 2530–2548.

[29] Gautam, A., Khan, Z.A., Gani, A., Asjad, M., 2025. Identification, ranking and prioritization of Key Performance Indicators for evaluating greenness of manufactured products. Green Technologies and Sustainability 3 (1), 100114.

[30] Lewandowska, A., Kurczewski, P., Joachimiak-Lechman, K., Zabłocki, M., 2021. Environmental Life Cycle Assessment of Refrigerator Modelled with Application of Various Electricity Mixes and Technologies. Energies 14 (17), 5350.

Biography

Nikolai Kelbel (*1995) studied industrial engineering at TU Darmstadt and has been a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2022. Nikolai Kelbel focuses on sustainable variant management and on the management of product complexity in industrial contexts.

Dr.-Ing. Alexander Keuper (*1991) studied mechanical engineering at RWTH Aachen University and has been a research assistant at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen since 2019, where he received his PhD and was appointed head of the Innovation Management Department in 2024. Dr.-Ing. Alexander Keuper works at the intersection of complexity management, development management, and data intelligence to support sustainable innovation under increasing system complexity.

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh (*1958) has the Chair of Production Engineering at RWTH Aachen University and is a member of the Board of Directors of the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen and the Fraunhofer Institute for Production Technology IPT in Aachen. As a scientist and entrepreneur, Prof. Günther Schuh focuses on disruptive innovations, information and production management, and sustainable mobility solutions.

8th Conference on Production Systems and Logistics

Procurement Target System of the Future

Theresa-Franziska Hinrichsen
Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany

Abstract

In recent years, procurement requirements have become more complex and dynamic. Companies are confronted with changing requirements due to factors such as politics and evolving consumer behavior. At the same time, they have become more vulnerable to disruptions. These changes have led to a paradigm shift in strategic procurement. Procurement is no longer viewed as merely a support function, but rather as a value-adding one. This creates a need to expand traditional targets, such as quality, quantity, location, and costs. To remain competitive, it is necessary to examine the understanding of targets in greater depth and to adapt existing target models. Currently, there is a lack of systematic analysis of future procurement targets. This paper aims to contribute to the strategic orientation of procurement. The goal is to identify traditional and expanded procurement targets and examine their interdependencies. The result is a target model that incorporates both traditional and expanded procurement targets.

Keywords

Procurement; Strategic Targets; Target System; Resilience; Order Processing

1. Introduction

The crises of recent years have led to a paradigm shift in strategic procurement [1]. Accordingly, strategic procurement is now considered a proactive, long-term integration of suppliers based on a strategic, value-creating understanding rather than a support function [2][3]. This shift underlines the need to expand traditional objectives [4]. In order to remain competitive, the underlying understanding of procurement objectives must be reassessed and existing objective models must be adapted. Currently, a systematic analysis of future procurement objectives and their interrelationships is lacking.

To address this gap, this paper examines which procurement objectives will be relevant in the future and how they relate to another. To answer this question, the paper first outlines the theoretical foundations for improved conceptual clarity. Chapter 2 discusses the state of the art in procurement objectives. This discussion is based on a literature review (Section 2.1) and an expert survey involving seven industry practitioners (Section 2.2). The expert survey assesses the future relevance of traditional procurement objectives and identifies additional objectives recognized in practice. Furthermore, existing procurement objective systems are analyzed to determine the extent to which they reflect this expanded understanding of objectives. Chapter 3 addresses the identified gap by presenting an appropriate target system that incorporates both traditional and extended procurement objectives. The paper concludes with a summary and outlook.

1.1 Definition of procurement

The task of strategic procurement includes defining the strategy and framework conditions for shaping procurement activities, as well as risk and supplier management to create potential for success. These

DOI: https://doi.org/10.15488/21015
ISSN: 2701-6277

decisions ensure the survival of the company by maintaining competitiveness and securing its market position. The choice of procurement strategy has a direct influence on the corporate and competitive strategy of supply processes [2]. Strategic procurement forms the orientation framework for operational procurement [1]. The focus of procurement is on the procurement object. These are all material goods and services that are procured [5]. Using the SCOR model, procurement can be classified within the supply chain. In the supply chain, procurement is located at the interface between suppliers and the company itself. Strategic procurement therefore has a strong influence on the company's innovative capacity, as it represents the interface with the supplier as an innovation partner. Through the optimal design of long-term, partnership-based supplier-customer relationships, supported by strategic supplier management, the procurement function becomes an integral part of the supply chain [6].

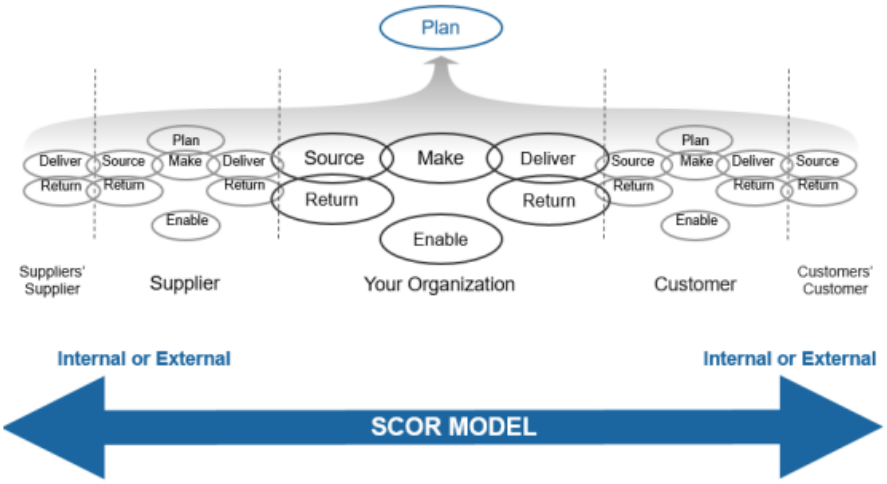


Figure 1: SCOR Model [6]

1.2 Target definition and target system

"Objectives form the basis for evaluating success or failure […]" [7]. They form the basis for corporate actions and target decisions. An objective is a possible circumstance (e.g., conditions, actions, relationships) that is desired to be achieved [8]. In the context of an action system, several interrelated goals are considered. When the general system model is applied, a goal system becomes a set of goals and relationships between them. A goal system represents linguistically formulated ideas of people, organizations, or societies [8]. A target system must meet two requirements: requirements at the target level and requirements at the target hierarchy level. In the context of this paper, the requirement at the target hierarchy level must be given special consideration. Typically, there are two challenges in decision-making: conflicts between targets and incomplete information. This requires a clear direction for simultaneously pursued targets in a target system, based on a ranking. Compatibility means that each element within the target system supports the system's primary goal [9]. There are three types of causal relationships [10]:

- In the case of an indifferent relationship, the implementation of one objective has no effect on the implementation of another objective.
- In the case of a complementary relationship, the implementation of one objective supports the realization of another objective.
- In the case of a competing relationship, the realization of one objective disables or prevents the realization of another objective.

2. State of the art procurement targets

2.1 Traditional targets

The primary objective of strategic procurement is to ensure long-term supply, reduce costs, and improve quality [11]. Other objectives mentioned in the literature include date, quantity, and location. The cost objective includes all costs incurred in procurement, e.g., price, logistics costs. The date objective covers the correspondence between the ordered and actual delivery date of the procurement item. The quantity objective describes the correspondence between the quantity delivered and the quantity ordered. The location objective describes the correspondence between the specified and actual place of performance [1]. The quality objective describes the correspondence between the product or service specification and the defined quality [12].

These objectives represent the traditional procurement goals. The paradigm shift in purchasing in recent years has led to an expansion of procurement objectives. According to the literature, these objectives include resilience, innovation, and sustainability. However, the term resilience is not uniformly defined in the literature. For the purposes of this paper, the term is defined as the ability to change. The goal of sustainability consists of three dimensions: economic, ecological, and social. The focus of economic sustainability is on sustainable management, prioritizing the well-being of society and the environment over profit maximization. The focus of ecological sustainability is on the environment. Social sustainability focuses on people and society [13]. The goal of sustainability is to bring the environment, economy, and people/society into balance, whereby the dimensions are in conflict with each other [14]. The goal of innovation means the targeted advancement of new innovations by adopting the technological expertise and ideas of suppliers [15]. From the perspective of active, value-creating, strategic procurement, as well as the expansion of goals to include sustainability, the goal of cost must be viewed in a broader context [16,4]. The Total Cost of Ownership (TCO) approach considers total costs and thus expands on the pure consideration of procurement costs [17]. The approach takes into account all costs throughout the life cycle, which are divided into three categories: procurement, use, and disposal. In addition to purely economic aspects, environmental factors and sustainability criteria also play a role in the TCO analysis [18]. This analysis is only found to a limited extent in literature. Until now, the classification has been divided into four categories [5]: direct material costs, inventory costs, order processing costs, and shortage costs.

2.2 Analysis of future targets

Seven companies were surveyed to analyze future targets. The companies surveyed were from the medical technology, electronics, metalworking, and automotive industries. The survey examined the future relevance of traditional procurement goals on the one hand and the complementary goals that experts see in practice on the other. The underlying hypotheses are as follows: 1. Traditional procurement goals will remain relevant in the future 2. Practice expands procurement goals 3. Traditional procurement goals are and will continue to be prioritized over expanded procurement goals in the future.

The first hypothesis can be confirmed. The majority of experts continue to attribute high to very high relevance to traditional procurement objectives. The procurement goal of quality will be considered more relevant in the future, while the goal of cost will be considered less relevant. While 100% of experts currently consider the goal of cost to be “highly relevant,” approximately 15% consider its future relevance to be moderate. The second hypothesis is refuted. The experts do not mention any further objectives. The third hypothesis cannot be clearly confirmed. However, the first part of the hypothesis can be confirmed, namely that traditional procurement objectives are currently rated higher overall than extended procurement objectives. The second part of the hypothesis has been refuted. It appears that the extended objectives are becoming more important.

2.3 Procurement target systems

The literature analysis shows that only a few authors classify procurement objectives within a target system. Most distinguish between strategic and operational objectives, including [19]. A further distinction is made between material and formal objectives, including [20] and customer and business objectives, including [21]. These approaches can be criticized for failing to consider influencing factors.

To close this gap, Schuh et al. [10] propose a target system approach based on Hess and Kleinlein [22] and Masing [23]. The authors distinguish between two dimensions: scope and target type, and two higher-level categories: non-profit and lifecycle [10]. In the scope dimension, a distinction is made between goals that relate to the process and those that relate to the object. In addition, a distinction is made between internal and cross-company goals. The goal type dimension considers the goals of quality and costs. At a higher level, the non-profit category encompasses all goals that have interactions between the four individual categories. Cost objectives that apply to objects, processes, and both within and across companies are assigned to the higher-level category of lifecycle [10]. One criticism of the target system is its focus on the two objectives of cost and quality. The authors assume that all objectives are directly related to the two objectives. However, there is no justification for this assumption.

Another approach comes from Erlach [24]. He refers to production goals, but takes goal relationships into account. Due to the interdependence of production and procurement, the transferability of this model to procurement will be examined below. He also divides his approach into four dimensions to which sub-goals can be assigned. These form the goal system based on their interrelationships. The author relates the objectives to the success factors that specify which customer requirements must be met by which performance characteristics in order for a company to be successful in the market [24]. The dimensions are: variability, quality, speed, and costs. There are four types of goal conflicts: contrary relationship, contradictory relationship, subordination relationship, and compatibility relationship [24]. It should be noted that, in contrast to the target system according to Schuh et al. [10], the dependencies between the dimensions are examined and presented in greater depth. Additionally, a more detailed distinction is made between the dimensions.

All in all, it should be noted that the target systems frequently mentioned in the literature do not reflect the interrelationship between the targets. The target system proposed by Schuh et al. [10] focuses on quality and costs in order to close this gap. Taking into account the procurement targets identified in this paper, this approach is not sufficient. A better approach is the target system according to Erlach [24] which is based on a broader level. It also depicts the interrelationships between the targets in a multidimensional way.

3. Procurement target system of the future

3.1 Dimensions

The weaknesses of the target system approaches discussed in Section 2.3 highlight the necessity of modeling a supplementary target system based on them. The procurement target system is based on the target system described by Erlach [24]. Due to its focus on production, this system must be adapted to procurement. This includes both a critical examination of the dimensions described in terms of their transferability or validity for procurement and in terms of necessary additions and target conflicts.

In terms of procurement, the variability dimension provides information about the ability to change, for example, by switching to a different supplier or using a different material. The quality dimension provides information about the quality required and delivered by the supplier. Speed is the time required for the delivery process. Cost-effectiveness refers to the relationship between performance and costs. Accordingly, the dimensions in the model can be applied to procurement. However, taking into account the identified traditional and extended objectives, these dimensions are not sufficient. Sustainability is only mentioned in a broader sense. According to the definition of sustainability, a distinction must be made between economic,

ecological, and social sustainability. Economic sustainability corresponds to the cost effectiveness dimension. Due to the increasing relevance of sustainability, the term economic sustainability is chosen for the dimension of cost effectiveness, the understanding of which goes beyond the consideration of economic efficiency and also takes social and ecological costs into account. Erlach [24] does not include this consideration. Social and ecological sustainability, on the other hand, cannot be assigned to any dimension. In accordance with the definition of resilience, the goal of resilience and the promotion of innovation are assigned to the variability dimension. The goals of quality, quantity, and location are assigned to the dimension of quality. This results in a target system with six dimensions.

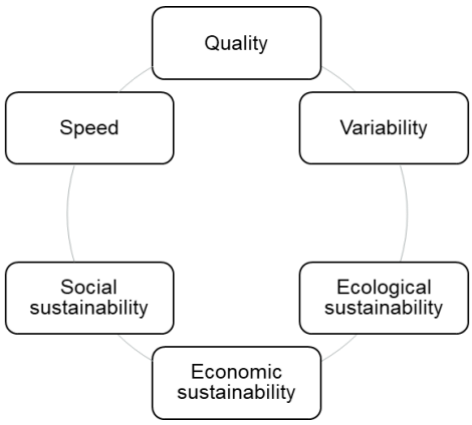


Figure 2: Target system of the future

3.2 Target relationships

The developed target system has three types of target relationships: complementary, indifferent, and competing. The complementary relationship corresponds to the target of compatibility. This means that improving a target in dimension A leads to an improvement in the target in dimension B. This also applies in the opposite direction. The indifferent relationship corresponds to goal subordination. The indifferent relationship states that the improvement or deterioration of a goal in dimension A has no effect on a goal in dimension B. This also applies in the opposite direction. The competing relationship summarizes contradictory and contrary contradictions. This relationship states that if a goal in dimension A deteriorates, this prevents or hinders the realization of the goal in dimension B. This also applies in the opposite direction.

The following section examines the conflicts of objectives between the dimensions in more detail using examples. The examples are for illustrative purposes only and are therefore mentioned in one direction. This has no effect on the direction of the relationship, which always exists in both directions.

3.2.1 Indifferent relationship

Variability and speed: indifferent relationship

In addition to the complementary relationship, there is an indifferent relationship between variability and speed. This means that a change in variability or speed has neither a positive nor a negative impact on the other dimension. For example, greater flexibility due to multiple suppliers does not necessarily lead to higher or lower speed.

Variability and economic sustainability: indifferent relationship

The indifferent relationship relates to variability and shortage costs, direct and indirect operating costs, spare parts costs, and disposal costs. The increase in variability relates to the procurement process at the beginning of the life cycle, so there is no impact on usage and disposal costs, as these occur at the end of the life cycle.

Quality and economic sustainability: indifferent relationship

This relationship relates to quality and inventory costs, as well as direct and indirect operating and disposal costs. The increase in quality relates to the procurement process at the beginning of the life cycle, so there is no connection between this dimension and disposal costs, as these occur at the end of the life cycle.

Quality and social sustainability: indifferent relationship

This target relationship states that changes in quality have no impact on social sustainability, and vice versa. For example, improving product quality does not change social compatibility or fair remuneration for employees.

Economic sustainability and social sustainability: indifferent relationship

This target relationship indicates that changes in social sustainability or shortage costs have no impact on the achievement of the other dimension. An increase in social sustainability does not lead to an increase or decrease in delivery failures.

Environmental sustainability and economic sustainability: indifferent relationship

The indifferent trade-off relates to environmental sustainability and storage costs. For example, choosing resource-efficient substances has no impact on storage options.

Environmental sustainability and social sustainability: indifferent relationship

This target relationship states that changes in environmental sustainability or social sustainability have no impact on the other dimension. For example, choosing sustainable raw materials does not change working conditions.

Economic sustainability and speed: indifferent relationship

The indifferent relationship exists between usage and disposal costs and speed. The objective relationship states that a change in one dimension has no effect on the other dimension.

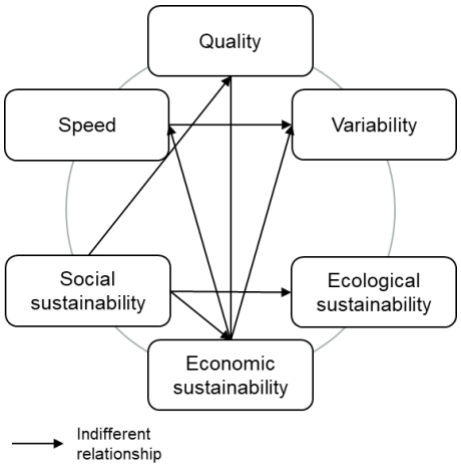


Figure 3: Indifferent relationship

3.2.2 Complementary relationship

Variability and speed: complementary relationshipThis conflict of objectives means that improving variability has a positive effect on speed, and improving speed has a positive effect on variability. For example, increasing flexibility can shorten delivery times by sourcing a procurement item from a supplier located closer to the customer.

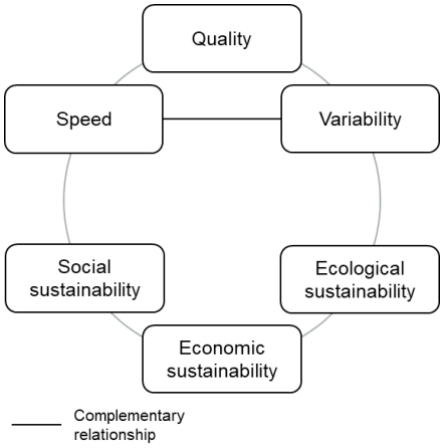


Figure 4: Complementary relationship

3.2.3 Competing relationship

Variability and economic sustainability: competing relationship

This target relationship indicates that improving variability makes it more difficult to maintain economic sustainability in terms of direct material, warehousing, and order processing costs. For example, greater flexibility requires multiple suppliers for material. This means that cost reduction potential through volume bundling cannot be realized.

Quality and speed: competing relationship

This target relationship states that changing the speed or quality makes it more difficult to achieve the target in the other dimension. For example, if a material is procured from a new supplier under time pressure, it may only be checked to a limited extent.

Economic sustainability and speed: competing relationship

This target relationship indicates that increasing speed makes economic sustainability more difficult to maintain in terms of spare parts, disposal costs, ordering costs, direct material costs, order processing costs, and shortage costs. For instance, delivery speed can be increased by using alternative, shorter means of transport with shorter transit times, such as air freight. However, shorter delivery times result in higher storage costs.

Quality and economic sustainability: competing relationship

This target relationship indicates that improving quality makes it more difficult to maintain economic sustainability. This conflict of objectives relates to ordering costs, excluding storage costs, and spare parts costs. For example, changing to higher-quality products can lead to higher direct material costs.

Quality and environmental sustainability: competing relationship

This target relationship indicates that increasing sustainability or quality makes it more difficult to achieve the target in the other dimension. For example, sustainable raw materials can influence product quality or properties.

Speed and economic sustainability: competing relationship

This target relationship indicates that increasing sustainability or speed makes it more difficult to achieve the target in the other dimension. This applies to warehousing and order processing costs. This may be the case, for example, if a sustainable means of transport is preferred to a less sustainable one, e.g., sea freight over air freight.

Economic sustainability and environmental sustainability: competing relationship

This target relationship indicates that increasing economic or environmental sustainability makes it more difficult to achieve the target in the other dimension. This applies to direct material costs, shortage costs, order processing costs, direct and indirect operating costs, spare parts costs, and disposal costs. For example, environmentally sustainable raw materials cost more.

Economic sustainability and social sustainability: competing relationship

This target relationship indicates that increasing economic or social sustainability makes it more difficult to achieve the target in the other dimension. Social working conditions in low-wage countries are still worse. This conflict applies in both directions. The target conflict relates to order processing costs, warehousing costs, direct material costs, direct and indirect operating costs, spare parts costs, and waste disposal costs.

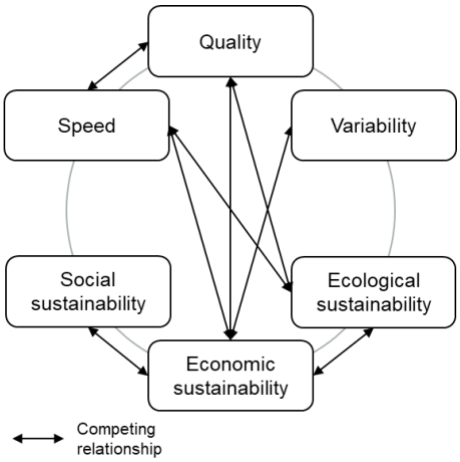


Figure 5: Competing relationship

4. Conclusion & outlook

This paper provides the basis for creating a more resilient strategic procurement process and aligning it with an increasingly complex and dynamic environment. Based on the literature review, it has been shown that traditional and expanded targets will become relevant in the future. Despite everything, it has become apparent that traditional targets are being expanded and the prioritization of goals is changing, which is particularly evident in the traditional procurement targets of cost and the expanded goal of resilience. Based on this, it can be concluded that companies are deviating from the trend of recent years of primarily procuring at minimal cost and are focusing on other targets. With regard to costs, the following applies: due to change, a pure cost analysis is insufficient. Until now, the majority of literature has distinguished between the categories of direct material, warehousing, order processing, and shortage costs. In line with the expansion of objectives and the transformation of procurement, these categories need to be broadened. The TCO approach is suitable for this purpose. Following this line of thinking, the consideration of cost items must also be expanded to include revenue items. Experts rated resilience as being just as relevant as costs in the future. This approach and its development should be observed and evaluated over the long term. Due to the increased number of crises and the impact of COVID-19 in recent years, the topic of resilience is particularly relevant in practice. Research into existing target systems has shown that they do not sufficiently consider the targets or their relationships. There are 17 target conflicts between the target dimensions, of which 1 is complementary, 8 are indifferent, and 8 are competing. Future work should examine how the new understanding of targets influences the design of strategic procurement and decision-making.

References

[1] Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K. (Eds.), 2008. Handbuch Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg.

[2] Lasch, R., 2014. Strategisches und operatives Logistikmanagement: Prozesse.

[3] Schuh, G. (Ed.), 2013. Einkaufsmanagement: Handbuch Produktion und Management 7, 2nd ed. ed. Springer, Dordrecht, 356 pp.

[4] Schuh, G., Hoppe, M., 2013. Einleitung und Abgrenzung, in: Schuh, G. (Ed.), Einkaufsmanagement. Handbuch Produktion und Management 7, 2nd ed. ed. Springer, Dordrecht, pp. 5–8.

[5] Beckmann, H., 2023. Grundkurs Beschaffungsmanagement. Springer Vieweg, Wiesbaden, Heidelberg, 406 pp.

[6] APICS, 2017. Supply Chain Operations Reference Model SCOR.

[7] Schuh, G., Hoppe, M., Schubert, J., Mangoldt, J. von, 2013. Lieferantenauswahl, in: Schuh, G. (Ed.), Einkaufsmanagement. Handbuch Produktion und Management 7, 2nd ed. ed. Springer, Dordrecht, pp. 183–254.

[8] Ropohl, G., 2009. Allgemeine Technologie: Eine Systemtheorie der Technik, 3., überarbeitete Auflage ed. Universitätsverlag Karlsruhe, Karlsruhe, 360 pp.

[9] Schmidt-Sudhoff, U., 1967. Unternehmerziele und unternehmerisches Zielsystem. Gabler Verlag, Wiesbaden, 169 pp.

[10] Schuh, G., Kreimeier, M., Essig, J., Schulte-Gehrmann, A.-L., 2013. Einkaufsstrategie, in: Schuh, G. (Ed.), Einkaufsmanagement. Handbuch Produktion und Management 7, 2nd ed. ed. Springer, Dordrecht, pp. 75–118.

[11] Grochla, E., Schönbohm, P., 1980. Beschaffung in der Unternehmung: Einführung in eine umfassende Beschaffungslehre. Poeschel, Stuttgart, 225 pp.

[12] Wiendahl, H.-H., 2011. Auftragsmanagement der industriellen Produktion: Grundlagen, Konfiguration, Einführung. Springer, Heidelberg, Dortrecht, 466 pp.

[13] Deutscher Bundestag (Ed.), 1998. Konzept Nachhaltigkeit: Vom Leitbild zur Umsetzung. Bundesanzeiger Verlagsgesellschaft mbH, Bonn.

[14] Kropp, A., 2019. Grundlagen der Nachhaltigen Entwicklung: Handlungsmöglichkeiten und Strategien zur Umsetzung, [1. Auflage] 2019 ed. Springer Gabler, Wiesbaden, Germany, 49 pp.

[15] Schuh, G., Guo, D., Hoppe, M., Ünlü, V., 2013. Steuerung der Lieferantenbasis, in: Schuh, G. (Ed.), Einkaufsmanagement. Handbuch Produktion und Management 7, 2nd ed. ed. Springer, Dordrecht, pp. 255–342.

[16] Braun, A.-T. Optimierungsmodell zur Planung der direkten Wiederwerwendung bei der Vermietung mobiler und langlebiger Investitionsgüter in Closed-loop supply Chains. Stuttgarter Beiträge zur Produktionsforschung. Fraunhofer Verlag.

[17] Schuh, G., Hacker, P.A., Wemhöner, H., 2013. Einkaufscontrolling, in: Schuh, G. (Ed.), Einkaufsmanagement. Handbuch Produktion und Management 7, 2nd ed. ed. Springer, Dordrecht, pp. 119–146.

[18] Konradin-Verlag Robert Kohlhammer GmbH, 2024. Total Cost of Ownership (TCO): So behält der Einkauf die Gesamtkosten im Blick. https://beschaffung-aktuell.industrie.de/einkaufscontrolling/total-cost-of-ownership-tco-gesamtkosten-einkauf/. Accessed 19 March 2025.

[19] Friedl, B., 1990. Grundlagen des Beschaffungscontrolling. Zugl.: Tübingen, Univ., Diss., 1988/89. Duncker u. Humblot, Berlin, 320 pp.

[20] Large, R., 2013. Strategisches Beschaffungsmanagement: Eine praxisorientierte Einführung ; mit Fallstudien, 5., vollst. überarb. Aufl. ed. Springer Gabler, Wiesbaden, 369 pp.

[21] Kuhn, A., 2008. Logistikprozesse in Industrie und Handel, in: Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K. (Eds.), Handbuch Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 215–253.

[22] Heß, G., Kleinlein, A.-C., 2021. Resilienz im Einkauf. Springer Fachmedien Wiesbaden, Wiesbaden, 194 pp.

[23] Masing, W., 2021. Masing Handbuch Qualitätsmanagement, 7., überarbeitete Auflage ed. Hanser, München, 1091 pp.

[24] Erlach, K., 2020. Wertstromdesign: Der Weg zur schlanken Fabrik, 3. Auflage ed. Springer Vieweg, Berlin, Heidelberg, 414 pp.

Biography

Theresa-Franziska Hinrichsen (*1994) studied Industrial Engineering and Business Administration in Hamburg and Mannheim. Since 2021, she has been a research fellow at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA with focus on production planning and control, digitalization and resilience.

8th Conference on Production Systems and Logistics

Forecasting Material Flow Composition in the German Returnable Beverage Container System

Leia Betting[1], Jakob Becker[1,2], Philipp Kempkes[3], Konrad Mielke[3], Markus Pauly[1,4]

[1]*Department of Statistics, TU Dortmund University, Dortmund, Germany*
[2]*Graduate School of Logistics, Dortmund, Germany*
[3]*Digitalisierung Anlagentechnik, REMONDIS SmartRec GmbH, Lünen, Germany*
[4]*Research Center Trustworthy Data Science and Security, UA Ruhr, Dortmund, Germany*

Abstract

Fluctuations in the volume of returned beverage containers can lead to challenges in resource allocation and workforce planning, particularly for smaller collection and counting centers. Forecasting of incoming beverage container quantities is therefore of interest. Accurate predictions enable more efficient distribution of personnel and operational resources across collection and counting facilities, thereby improving logistical performance and cost-effectiveness within the overall deposit-return system. This study presents the application and evaluation of data-driven forecasting models designed to predict the quantity of returned beverage containers differentiated by material type across five counting centers in Germany on a horizon of four weeks and one year. The modeling approach integrates both traditional time series techniques and machine learning algorithms. In addition to the data from the five counting centers, external variables such as meteorological conditions and seasonal and holiday information are incorporated into the analysis. To appropriately account for temporal dependencies during model assessment, a rolling cross-validation scheme is employed. Model performance is quantified using the Mean Squared Error metric. Our findings indicate that DHR (Dynamic Harmonic Regression) and TBATS (Trigonometric Box-Cox transform, ARMA errors, Trend and Seasonal components) models yield promising results for forecasting the return volumes of the different beverage container materials. Their effectiveness in our use-case can be attributed to their explicit capability to capture complex seasonal patterns inherent in the data. The incorporation of external explanatory variables has mixed outcomes on model performance.

Keywords

Time Series Prediction; Statistical Models; Machine Learning; Plastic; Recycling; Beverage Containers; Sustainability

1. Introduction

In Germany, the beverage container return system operates deposit-based. With a reported recycling rate of approximately 97.6% [1] – indicating that nearly all of the returned beverage containers could be reprocessed – the system demonstrates high operational efficiency. Nonetheless, certain structural bottlenecks persist within the system. A key consideration arises regarding the trade-off between achieving cleaner material streams through reduced material diversity in the stream and maintaining separate recycling chains for beverage containers, which require distinct collection infrastructures, processing plants, logistics networks and personnel resources.

DOI: https://doi.org/10.15488/21016
ISSN: 2701-6277

Enhancing the efficiency of this dedicated recycling chain necessitates accurate short-term forecasts of beverage container quantities arriving at counting centers. Reliable predictive information enables improved planning across several dimensions – processing, capacity allocation, workforce scheduling and logistical coordination – all contributing to reduced bottlenecks throughout subsequent stages in the recycling process. Such optimization not only improves material quality but also lowers operational costs by ensuring more efficient resource utilization. Furthermore, accurate forecasting supports strategic long-term decisions such as expanding existing facilities or constructing new sorting plants. Ultimately, these advancements may increase overall system efficiency and strengthen the economic competitiveness of recycled materials relative to virgin plastics, thereby contributing to Target 12.5 of the sustainable development goals set by the United Nations [2].

Based on our review of the available literature, we did not identify studies that specifically address the forecasting of beverage container volumes. Existing studies focus on topics such as waste generation, recycling rates or material flow analyses. Consequently, this study constitutes an initial attempt to address this topic. The absence of prior work can likely be attributed to different waste management systems across countries, as many do not recycle beverage containers separately but instead dispose of them together with household waste. Nevertheless, within the broader contexts of municipal solid waste and plastic packaging waste, forecasting has received increasing attention in recent years. For instance, [3] investigate several modeling approaches for municipal solid waste forecasting in the Czech Republic and identify random forest as yielding the highest predictive accuracy. Similarly, [4] develop a predictive model for post-consumer plastic packaging waste in the Netherlands, including multiple categories for beverage containers. Furthermore, [5] explore neural networks as a methodological approach for forecasting municipal waste generation in Croatia.

This research investigates the prediction of beverage container returns across different beverage container materials at five German beverage container counting centers with seven forecasting approaches: Naive method, SARIMA (Seasonal Auto Regressive Integrated Moving Average), DHR (Dynamic Harmonic Regression), TBATS (Trigonometric Box-Cox transform, ARMA errors, Trend and Seasonal components), RF (Random Forest), XGBoost (Extreme Gradient Boosting) and LSTMs (Long Short-Term Memory Networks). The study addresses three central research questions:

1. Can the utilized models outperform the naive method that uses the most recent observation as a forecast?
2. Does incorporating external factors improve the predictive performance of the models?
3. Does forecast performance deteriorate as the forecast horizon increases?

The external variables considered include lagged weather data, motivated by [6] and [7], alongside calendar-based characteristics such as the season, the weekday, school holidays, public holidays, the number of days to the previous public holiday and the number of days to the next public holiday. Model performance is evaluated using rolling time series cross-validation with Mean Squared Error (MSE) serving as the primary error metric.

The paper is organized as follows: Section 2 details the dataset utilized. Section 3 summarizes the statistical forecasting methodologies. Section 4 presents a discussion of the analysis as well as the results. Finally, Section 5 concludes with an evaluation of the top-performing models and outlines potential directions for future research.

2. Data description

2.1 General process

Plastic recycling involves four principal stages: (1) collection, (2) sorting, (3) cleaning and (4) reprocessing [2]. In Germany, the returnable beverage container sector – largely composed of plastic materials – is organized through a deposit-based system. Within this framework, consumers return used beverage containers to collection centers, which primarily serve to forward these containers to counting centers. Additionally, sorting occurs at the collection stage via reverse vending machines and the processed beverage containers are directly forwarded to the reprocessing facilities.

2.2 Set-up

Five counting centers located in Germany constitute the primary source of data for this study. At each facility, returned beverage containers are categorized according to material type: *aluminum* (ALU), *thick-walled ALU*, *glass*, *high-density polyethylene* (HDPE), *low-density polyethylene* (LDPE), *polyethylene terephthalate* (PET), *PET/metals*, *PET with nylon* (PETN), *polypropylene* (PP), *other plastics*, *steel* and *unknown (UNK)*. Due to low occurrence numbers in certain categories, materials were summarized as follows: *ALU* and *thick-walled ALU* into a single category (ALU) and *HDPE*, *LDPE*, *PET/metals*, *PP* and *other plastics* into one aggregated group (plastics). The *unknown (UNK)* category represents beverage containers that could not be classified under any existing material class. A list of all materials and its abbreviations can be seen in Table 1.

The dataset comprises both the **daily number of containers** and their corresponding **total weight in tons** across all material types and centers. For the period between January 2nd, 2023 and December 29th, 2023, data are disaggregated by center, whereas from January 2nd, 2020 to December 31st, 2022 only aggregated data across all five centers are available. In total, observations span four years, from 2020 through 2023.

2.3 Data preprocessing

Data preprocessing was conducted in three sequential steps. First, as three of the four years contained only aggregated data across all centers, data for 2023 were likewise aggregated for consistency. Second, observations recorded on weekends or nationwide public holidays were excluded because operational downtime resulted in near-zero values that could distort time-series analyses without providing meaningful information. Third, the daily counts of returned containers were normalized using min-max normalization following [8] to ensure comparability among different material categories.

After applying these procedures, the resulting time series consisted of 1004 observations over four years. This corresponds to an annual seasonal period of 1004/4 = 251 observations.

2.4 External data

To account for potential external influences on container return rates or beverage consumption patterns, several external covariates were incorporated. Meteorological data – including **precipitation**, **temperature** and **sunshine duration** averaged over a monthly period – were obtained from the German Weather Service [9] and interpolated to daily resolution using cubic spline interpolation as described by [10]. Furthermore, information regarding **public holidays** and **school holidays** were integrated and weighted by the share of population of the respective federal states compared to the total population of Germany. Population figures correspond to those reported by the Federal Statistical Office of Germany as of December 31st, 2019 [11] and were held constant throughout the observation period. Additionally, other calendar-based features were included: **weekday**, taking values *Monday* through *Friday* and **season**, taking values *spring*, *summer*, *autumn*, or *winter*, corresponding to the meteorological season [12].

3. Methodology

The applied forecasting approaches can be categorized into two overarching groups: statistical time series methods and machine learning (ML) models. The statistical methods include the naive method, SARIMA, DHR and TBATS, whereas the ML-based techniques comprise two state-of-the-art tree-based ensemble methods as well as LSTM networks. The methods were chosen on the basis of [13]. Moreover, DHR and TBATS are included due to their capability of modeling complex seasonality.

The **naive method** relies solely on the most recent observation, using its value directly as a forecast for the subsequent time steps. As outlined by [14], this simple approach often yields satisfactory results in various applications and therefore serves as a benchmark for evaluating more complex forecasting techniques.

The **Seasonal Auto Regressive Integrated Moving Average (SARIMA)** model integrates concepts of autoregression, differencing to achieve stationarity and moving averages within a unified framework. The in-depth theoretical foundation is provided by [15]. SARIMA has been widely adopted in forecasting tasks, including applications related to waste forecasting [16].

The **Dynamic Harmonic Regression (DHR)** extends traditional regression by incorporating Fourier terms as predictors [17], allowing the seasonality to have any length and to even incorporate multiple seasonal patterns [14]. In contrast to SARIMA, seasonal patterns are thus modeled explicitly rather than through seasonal differencing.

The **Trigonometric Box-Cox transform, ARMA errors, Trend and Seasonal components (TBATS)** model combines trend estimation, seasonality representation through trigonometric functions based on Fourier terms and ARMA error modeling within an additive structure. To stabilize variance in the data, a Box-Cox transformation may be applied prior to modeling [18].

The **Random Forest (RF)** algorithm constructs an ensemble of decision trees trained on bootstrap samples with randomly selected subsets of features at each split [19]. By aggregating predictions from decorrelated trees, RF achieves good performance on tabular datasets [20] but also in time series applications [21].

Similarly, **eXtreme Gradient Boosting (XGBoost)** represents an advanced variant of gradient boosting that sequentially builds weak learners – typically shallow trees – to iteratively improve overall predictive accuracy through weighting mechanisms. XGBoost has shown strong performance in waste prediction scenarios, see e.g. [22].

Finally, **Long Short-Term Memory Networks (LSTMs)** constitute a specialized form of recurrent neural networks (RNNs) designed to process sequential data while mitigating the vanishing gradient problem inherent in traditional RNNs. Through their internal gating mechanisms – comprising input, output and forget gates – LSTMs maintain both short-term and long-term dependencies efficiently. Consequently, they have become prominent tools for modeling temporal patterns across diverse application areas including waste generation forecasting [23].

Overall, it should be noted that while the naive method, SARIMA and TBATS are limited to univariate time series settings without external covariates, methods such as DHR, RF, XGBoost and LSTMs are capable of integrating external information sources to potentially enhance predictive performance.

4. Analysis

4.1 Setup

The objective of this analysis is to identify the most suitable forecasting method for modeling the data based on the **Mean Squared Error (MSE)** as the primary metric for predictive performance.

According to [14], MSE quantifies the average squared deviation between predicted and observed values, while its rooted form – the Root MSE (RMSE) – expresses model accuracy on the same scale as the original data. In this study, however, RMSE is not utilized since all variables have been normalized, rendering the scale as irrelevant.

Model performance is assessed using **rolling Time Series Cross-Validation (TSCV)**. Unlike standard cross-validation procedures, TSCV ensures that all training observations precede those used for validation, thereby preventing temporal information leakage – a common issue when dealing with autocorrelated data. The number of folds in the validation process is determined by specifying an *initial window* and a *forecast horizon*. Initially, a model is trained using observations within the initial window and employed to predict values over the subsequent forecast horizon. The resulting predictions are evaluated via MSE. The window is then advanced by one observation – hence the TSCV is "rolling" – and prediction and evaluation steps are repeated. This procedure is performed until further rolling would exceed dataset boundaries. This procedure is graphically depicted in Figure 1 adapted from [14].

When incorporating **external covariates** into forecasting models, it is crucial that these predictors are available across the entire forecast horizon. The calendar-based features pose no challenge in this regard; however, weather-related information require backward shifting for availability on the full forecast horizon. This is also more realistic as beverage container returns are more likely to be influenced by weather conditions prevailing during its consumption and not during its arrival at the counting center. A cross-correlation analysis between returned beverage container quantities and weather variables indicated that applying a lag of 20 working days provides a compromise across different materials and meteorological variables. This choice reflects a good tradeoff between varying optimal lags among distinct material types and weather indicators and additionally allows for forecasting 20 working days ahead.

A descriptive examination of material composition on the dataset revealed that approximately 60% of returned containers were classified as PET, around 30% as ALU, roughly 7% as steel, about 2% as UNK, while the remaining 1% were distributed among minor materials. Consequently, PET constitutes the primary focus of subsequent analyses due to its dominant share in total return volumes.

4.2 Implementation

The analysis was conducted using the statistical programming software *R* [24]. The corresponding source code utilized for the analysis is publicly accessible via GitHub at https://github.com/jakobTU/plastic_bottle_forecasting.

For the training of ML models, each material's time series was converted into a tabular dataset. To enhance model performance for RF and XGBoost, the data was detrended by fitting a linear regression model [25] and the residuals from this model were subsequently employed as input data. This procedure mitigates limitations inherent to tree-based algorithms, which are constrained to the range of observed values during training and therefore unable to extrapolate trends effectively.

Seasonality within the ML models was addressed through Fourier transformations implemented via the *caretForecast* package, analogous to the modeling approach applied in DHR. For each time series, forecasting new observations involved designating the subsequent value as the target variable while utilizing the preceding 20 days' values as covariates.

A nested TSCV procedure was embedded within the broader TSCV framework outlined in Section 4.1 to determine suitable hyperparameter configurations for RF and XGBoost. The specific hyperparameter combinations examined are provided in Table 2 and 3. All possible hyperparameter combinations were systematically evaluated and those yielding the lowest RMSE across a four-day forecast horizon were selected for the respective fold. The model corresponding to this best hyperparameter combination was subsequently used to generate forecasts for the validation set horizon of the outer TSCV, which were

evaluated using the MSE. For the LSTM, the only architecture considered consisted of two layers with 50 neurons each, followed by dropout layers with a dropout rate of 0.2 after both hidden layers. Hence no hyperparameter tuning was necessary and the nested TSCV was omitted.

For the outer TSCV implementation, an initial window comprising 50% of available data – corresponding to 502 observations – was employed. Forecast horizons were set to 20 days for short-term predictions and 251 days for long-term forecasts.

4.3 Short-term forecasts without external information

The performance of the seven forecasting methods was evaluated over forecast horizons ranging from 1 to 20 days ahead using the previously described TSCV procedure without taking external covariates into account.

Figure 2 illustrates the MSE across the different forecast horizons for the normalized quantity of PET containers. Among all models, TBATS demonstrates superior performance, consistently achieving the lowest MSE values across all horizons. Up to a forecast horizon of ten days, TBATS exhibits a gradual increase in MSE – an expected outcome due to the growing uncertainty associated with forecasts further into the future – after which the MSE stabilizes around 0.0084.

The RF also yields relatively low MSE values overall. However, its error increases more sharply with expanding forecast horizons compared to TBATS. XGBoost and the LSTM display similar progression patterns to RF but with steeper increases in MSE. DHR shows a marked rise in MSE up to ten days, followed by a decline that results in the second-best performance at a horizon of twenty days. In contrast, both the naive method and SARIMA perform considerably worse across all horizons, with SARIMA exhibiting particularly poor results up to fifteen days.

Table 4 provides an overview of the mean MSE values across the twenty forecast horizons for all materials and methods. The results from Table 4 indicate that maintaining an unchanged forecast for twenty consecutive working days leads to a substantial deterioration in predictive performance, as the naive method is outperformed by all other methods across most materials except for SARIMA on PET. While SARIMA generally does not demonstrate exceptional predictive performances – it ranks first only for glass where the LSTM follows closely – DHR achieves the best performances for four of the seven materials and nearly matches the LSTM's top performance on PETN. However, DHR shows suboptimal performances for glass and PET. TBATS performs comparably well: it nearly equals DHR's performance on DHR's four best-performing materials, slightly underperforms on PETN, yet surpasses DHR substantially for glass and especially PET where its forecasting performance is superior to all other considered methods. RF demonstrates solid but less competitive results overall, slightly outperforming DHR only on glass and PET while lagging behind otherwise. XGBoost exhibits consistently inferior performance across all materials examined. The LSTM delivers strong results for glass and particularly for PETN where it ranks first but performs less accurately on the other materials.

In conclusion, these findings reveal that advanced statistical procedures – particularly DHR, TBATS and RF – achieve accurate forecasting performance overall. Moreover, the naive method maintaining a constant forecast is outperformed by a large margin in almost all cases.

4.4 Short-term forecasts with external information

In this scenario, the external covariates described in Section 2.4 were integrated into model training, while evaluation across forecast horizons from 1 to 20 days was conducted using TSCV.

Figure 3 displays the resulting MSE values for PET containers when incorporating external covariates. Only DHR, RF, XGBoost and the LSTM support external covariates, but due to computational constraints the LSTM was omitted. TBATS without covariates is shown as a reference benchmark.

Across all cases examined, models including covariates achieve lower MSEs than their counterparts without covariates. Nevertheless, TBATS continues to yield the smallest errors throughout all forecast horizons. This finding suggests that while external covariates enhance predictive performance within individual models, they do not surpass TBATS performance when forecasting normalized PET quantities.

Among models utilizing covariates, RF achieves superior results relative to XGBoost and DHR across nearly all horizons except at a horizon of 1 where DHR slightly outperforms RF. For forecast up to ten days, DHR and XGBoost exhibit comparable behavior, beyond this range XGBoost's error increases more markedly whereas DHR remains relatively stable. At a twenty-day horizon DHR performs only marginally worse than RF. Comparing DHR variants with and without covariates reveals noticeable improvements for shorter horizons when external covariates are included. However, differences diminish toward longer forecast horizons between eighteen and twenty days.

Table 5 summarizes the mean MSE values averaged across the twenty forecast horizons per method and material. Generally, no massive performance gains are visible as DHR is still the best-performing method on more than half of the materials, for the other two materials RF performs best. Comparisons with Table 4 indicate that the inclusion of covariates leads to mixed outcomes. DHR with covariates produces lower errors than its counterpart without covariates for ALU, PETN, PET and UNK while performing worse on the other materials. Notably, for ALU and UNK, DHR with covariates surpasses the previous best-performing model without external covariates – although the improvement is marginal for UNK – demonstrating potential benefits from incorporating external information under specific conditions. This is supported by the performances of the RF with covariates which surpasses the performance of the RF without covariates for six of the seven materials, only for steel the performance worsens. On the contrary, for XGBoost only for plastics and PET marginal improvements can be noted while for the other five materials the inclusion had the opposite effect. However, it can be stated that for the predominant material PET all three methods incorporating covariates outperform its counterpart not taking external information into account. Regarding an overall improvement of the best-performing method for each material, the inclusion of external covariates only leads to an improvement for ALU and UNK indicating that overall the used external covariates do not have excelling benefit on performance.

4.5 Long-term forecasts

To assess long-term predictive capabilities, the presented TSCV procedure was applied to an extended forecast horizon of 251 working days and the MSE was calculated across the TSCV folds. In this scenario, only the previously identified high-performing models – DHR, TBATS, RF and LSTM – were considered. Again, the naive method served as a baseline. External covariates were incorporated into DHR and RF, whereas TBATS and the LSTM were applied without covariates due to the aforementioned methodological and computational constraints. The selection of external covariates was limited to the calendar-based information because in this scenario the covariates must be known one year in advance. For the weather-based covariates, future observations cannot be determined reliably for such a forecasting horizon.

Table 6 presents the MSE values broken down by material and forecasting method for the forecast horizon of 251 working days. The naive method achieves the lowest MSE for PETN and UNK, while performing worst for the remaining five materials. For PETN in particular, the naive approach outperforms all competing methods by a considerable margin, indicating that more complex models fail to improve upon the simple baseline for this material. This can be attributed to a positive trend of PETN until mid 2023 which, however, appears to be constant afterwards. DHR attains the best performance for PET and Steel and demonstrates consistently competitive results for most other materials. With the exception of PETN, where its performance deteriorates markedly, DHR remains close to the respective best-performing method across materials. TBATS exhibits performance pattern similar to DHR. It achieves the lowest MSE for Glass and delivers results comparable to DHR for most materials. However, for PETN its performance declines substantially, resulting in a markedly higher MSE relative to the naive benchmark. RF ranks first for ALU and Plastics

and performs comparably to the best method for Glass and PET. In contrast, its predictive performance is considerably lower for Steel and UNK. For PETN, RF ranks second but remains substantially inferior to the naive method. The LSTM achieves MSE values close to the best-performing method for Glass and Plastics but exhibits noticeably weaker performance for the remaining materials. Notably, it does not attain the lowest MSE for any material considered.

These results indicate that long-term forecasting in this context poses challenges to accurate forecasting. In particular, for PETN and UNK, the naive approach based on the corresponding observation from the previous year is advantageous over more complex models. Conversely, forecasts produced by DHR, TBATS and RF successfully capture general seasonal patterns and provide meaningful projections one year ahead as the MSE values of Table 6 are not too different from the values of Tables 4 and 5 suggesting that predictive performance declines over longer horizons, but to a lesser extent than might be expected.

5. Conclusion

In this study, seven time series and ML approaches were compared for modeling and forecasting the normalized quantities of seven beverage container materials processed across five German beverage container counting centers. Predictive performance was evaluated using a rolling TSCV with MSE serving as the primary evaluation metric. Furthermore, the analysis examined the effects of incorporating external covariates and extended forecasting over a long-term horizon of one year. Following the analysis, the research questions were systematically addressed:

1. Across all examined cases, both model types – with and without external covariates – outperform the naive method on a forecast horizon of twenty days.
2. The benefit of including external covariates is model- and material-dependent rather than universally advantageous. For the predominant material PET, all three models incorporating external covariates achieve improved performance compared to their counterparts without covariates. However, none exceeds the best-performing model without covariates for this material.
3. Forecasting accuracy declines as the forecast horizons increase, a pattern observed for both short-term monthly predictions and long-term annual forecasts. Despite this deterioration, even at a one-year horizon the performance levels achieved by the advanced statistical methods remained impressive, underscoring their robustness under extended temporal conditions.

Nevertheless, several avenues for improvement exist. Access to more granular data at the level of individual counting or collection centers could enable further studies to incorporate spatial information such as location-specific weather information or regional holiday schedules. Additionally, integrating socio-economic indicators – including income levels or demographic age structures – might help explain regional variations in return behavior. Future extensions of the analysis may include modifications to the preprocessing procedure, such as using the original values instead of min-max normalized series as forecasting targets, addressing outliers through removal or imputation and applying seasonal adjustments to individual materials. Additionally, employing alternative error metrics could enhance interpretability and modeling the aggregate volume of returned beverage containers may improve forecasts at the material-specific level.

Notably, as an outcome of this work, extensions of the well-performing methods identified herein are intended to be used in the near future for staff planning at the counting centers.

Acknowledgements

This work has been partly supported by the Research Center Trustworthy Data Science and Security (https://rc-trust.ai), one of the Research Alliance centers within the UARuhr (https://uaruhr.de). During the preparation of this paper, the authors used OpenAI's ChatGPT model to suggest minor language edits to improve readability. After using this tool, the authors have reviewed and revised the content as needed and take full responsibility for the content of the publication.

Appendix

Table 1: Overview of all materials in the data, including their abbreviations and assigned material categories used in the analysis.

Material	Abbreviation	Material category
Aluminum	ALU	ALU
Thick-walled aluminum	Thick-walled ALU	ALU
Glass	No abbreviation	Glass
High-density polyethylene	HDPE	Plastics
Low-density polyethylene	LDPE	Plastics
Polyethylene terephthalate	PET	PET
Polyethylene terephthalate/metals	PET/metals	Plastics
Polyethylene terephthalate with nylon	PETN	PETN
Polypropylene	PP	Plastics
Other plastics	No abbreviation	Plastics
Steel	No abbreviation	Steel
Unknown	No abbreviation	UNK

Table 2: Hyperparameter grid for the Random Forest model. All combinations of the listed values were evaluated. For an explanation of the hyperparameters we refer to [26].

Hyperparameter	Random Forest
mtry	4, 7, 10
splitrule	'variance'
min.node.size	1, 5, 10

Table 3: Hyperparameter grid for the Extreme Gradient Boosting model. All combinations of the listed values were evaluated. For an explanation of the hyperparameters we refer to [27].

Hyperparameter	Extreme Gradient Boosting
nrounds	100, 1000
max_depth	1, 5
eta	0.01, 0.3
gamma	0
colsample_bytree	1
min_child_weight	1
subsample	0.5, 0.8

Table 4: The arithmetic mean of the mean squared error (MSE) values across the 20 forecast horizons for each material and forecasting method without external covariates. The best-performing method per material, indicated by the lowest unrounded MSE, is highlighted in **bold**.

Material	Naive	SARIMA	DHR	TBATS	RF	XGBoost	LSTM
ALU	0.0152	0.0114	**0.0083**	0.0083	0.0088	0.0097	0.0106
Glass	0.0362	**0.0189**	0.0246	0.0206	0.0228	0.0295	0.0199
Plastics	0.0082	0.0049	**0.0048**	0.0049	0.0061	0.0064	0.0071
PETN	0.0183	0.0131	0.0128	0.0136	0.0149	0.0171	**0.0126**
PET	0.0127	0.0139	0.0097	**0.0078**	0.0092	0.0108	0.0099
Steel	0.0160	0.0111	**0.0097**	0.0099	0.0133	0.0143	0.0118
UNK	0.0230	0.0139	**0.0128**	0.0131	0.0162	0.0159	0.0144

Table 5: The arithmetic mean of the mean squared error (MSE) values across the 20 forecast horizons for each material and forecasting method including external covariates. The best-performing method per material, indicated by the lowest unrounded MSE, is highlighted in **bold**. Values marked with a dagger (†) indicate an improvement over the corresponding model without external covariates (see Table 4). The final column indicates whether the overall best-performing method for the respective material (considering Tables 4 and 5) incorporates external covariates.

Material	DHR	RF	XGBoost	Overall improvement?
ALU	**0.0077**†	0.0082†	0.0104	Yes
Glass	0.0265	**0.0227**†	0.0302	No
Plastics	**0.0048**	0.0054†	0.0063†	No
PETN	**0.0128**†	0.0147†	0.0173	No
PET	0.0091†	**0.0085**†	0.0103†	No
Steel	**0.0101**	0.0141	0.0149	No
UNK	**0.0128**†	0.0156†	0.0162	Yes

Table 6: The arithmetic mean of the mean squared error (MSE) values across the 251 forecast horizons for each material and the best-performing methods identified in Tables 4 and 5. Methods originating from Table 5 include external covariates, with the weather-related covariates excluded due to practical constraints. The best-performing method per material, indicated by the lowest unrounded MSE, is highlighted in **bold**.

Material	Naive	DHR	TBATS	RF	LSTM
ALU	0.0461	0.0100	0.0106	**0.0091**	0.0254
Glass	0.0607	0.0289	**0.0210**	0.0251	0.0211
Plastics	0.0238	0.0086	0.0095	**0.0082**	0.0098
PETN	**0.0167**	0.0297	0.1076	0.0285	0.7240
PET	0.0432	**0.0104**	0.0108	0.0134	0.0185
Steel	0.0285	**0.0129**	0.0159	0.0225	0.0251
UNK	**0.0162**	0.0194	0.0199	0.0277	0.0429

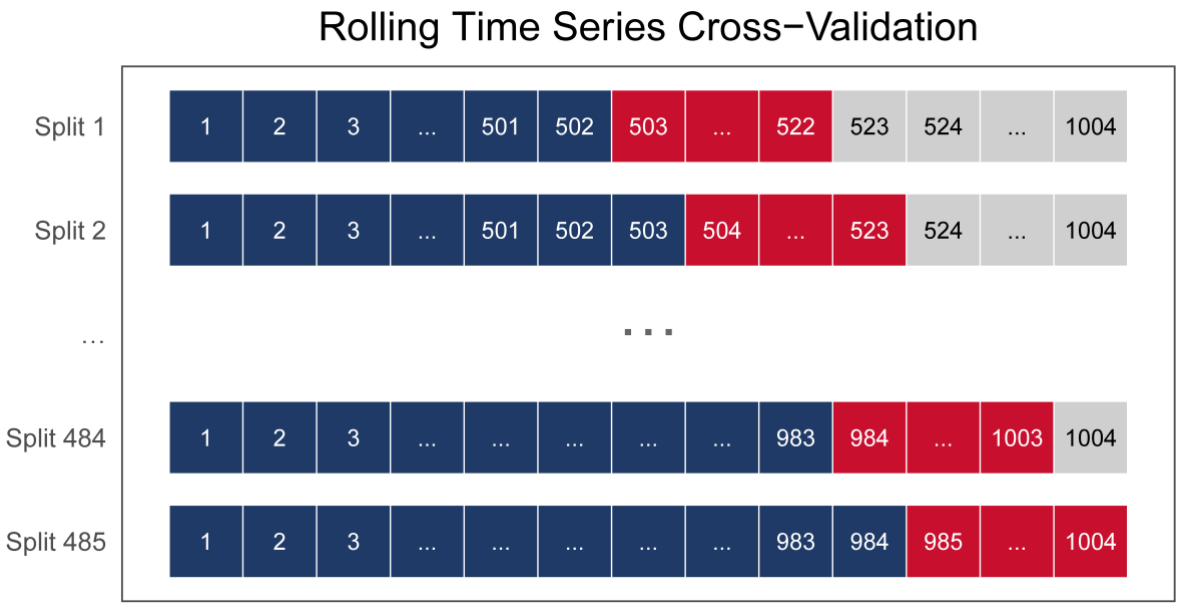


Figure 1: Schematic illustration of the rolling time series cross-validation configuration employed in this study. Blue blocks represent the training observations, red blocks denote the validation observations and gray blocks indicate future observations not used in the respective fold. This figure is adapted from [14].

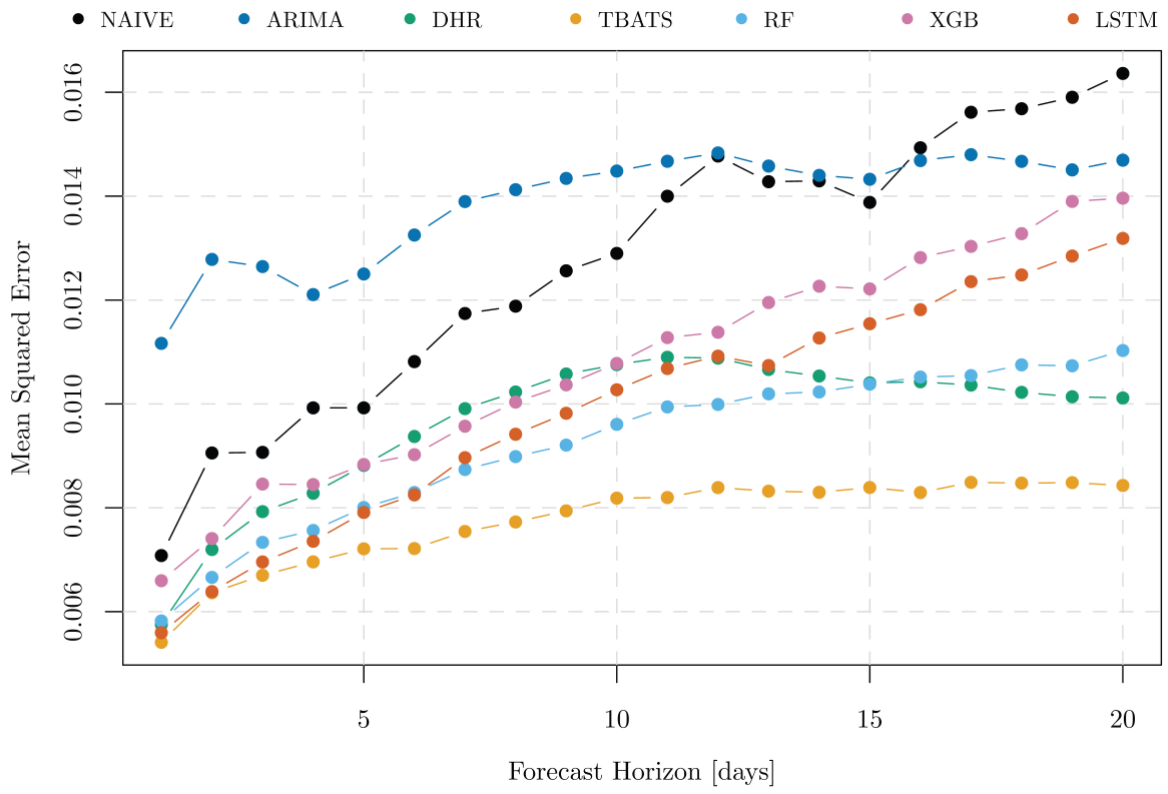


Figure 2: Development of the mean TSCV MSE values for the normalized quantity of PET containers for the different models. The displayed models do not incorporate external covariates.

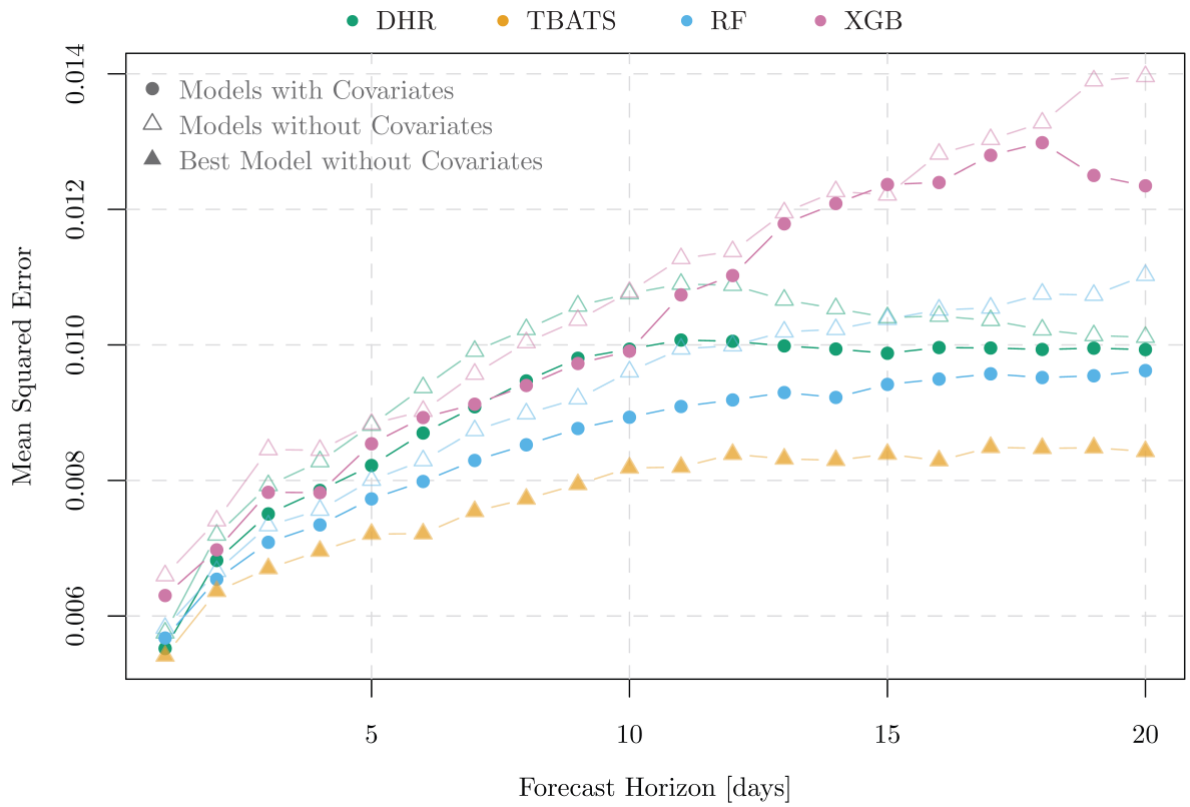


Figure 3: Development of the mean TSCV MSE values for the normalized quantity of PET containers for the different models capable of incorporating external covariates, compared to their counterparts without covariates. The TBATS model without covariates serves as a baseline reference.

References

[1] PETpla.net, 2024. German PET beverage bottle market 2023: more recycled content, less material use, high recycling rates. https://www.petpla.net/2024/11/12/german-pet-beverage-bottle-market-2023-more-recycled-content-less-material-use-high-recycling-rates/. Accessed 30 March 2026.

[2] United Nations Development Programme, 2023. Why aren't we recycling more plastic? https://stories.undp.org/why-arent-we-recycling-more-plastic. Accessed 30 March 2026.

[3] Rosecký, M., Šomplák, R., Slavík, J., Kalina, J., Bulková, G., Bednář, J., 2021. Predictive modelling as a tool for effective municipal waste management policy at different territorial levels. Journal of Environmental Management 291, 112584.

[4] Brouwer, M.T., van Velzen, Eggo U. Thoden, Augustinus, A., Soethoudt, H., Meester, S.d., Ragaert, K., 2018. Predictive model for the Dutch post-consumer plastic packaging recycling system and implications for the circular economy. Waste Management 71, 62–85.

[5] Ribic, B., Pezo, L., Sincic, D., Loncar, B., Voca, N., 2019. Predictive model for municipal waste generation using artificial neural networks - Case study City of Zagreb, Croatia. International Journal of Energy Research.

[6] Maimunah, Buliali, J.L., Saikhu, A., 2023. The use of LSTM model with lagged daily inputs for waste disposal prediction, in: 2023 Eighth International Conference on Informatics and Computing (ICIC), pp. 1–6.

[7] Vu, H.L., Ng, K.T.W., Bolingbroke, D., 2019. Time-lagged effects of weekly climatic and socio-economic factors on ANN municipal yard waste prediction models. Waste Management 84, 129–140.

[8] Han, J., Kamber, M., Pei, J., 2012. Data mining: concepts and techniques, 3rd ed. Morgan Kaufmann.

[9] Deutscher Wetterdienst, 2025. Monthly regional averages of climate variables in Germany. https://opendata.dwd.de/climate_environment/CDC/regional_averages_DE/monthly/. Accessed 30 March 2026.

[10] Burden, R.L., Faires, J.D., 2010. Numerical analysis, 9th ed. Brooks/Cole, Cengage Learning, Boston, MA.

[11] Statistisches Bundesamt, 2021. Bevölkerung: Fortschreibung des Bevölkerungsstandes – 2019. https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00134556. Accessed 30 March 2026.

[12] National Centers for Environmental Information (NCEI), 2016. Meteorological versus astronomical seasons. National Oceanic and Atmospheric Administration (NOAA). https://www.ncei.noaa.gov/news/meteorological-versus-astronomical-seasons. Accessed 30 March 2026.

[13] Becker, J., Kempkes, P., Mielke, K., Fendel, A., Bommert, A., Pauly, M. Forecasting plastic waste fractions: a recycling perspective, in: Proceedings of the Conference on Production Systems and Logistics, vol. 7, pp. 494–508.

[14] Hyndman, R.J., Athanasopoulos, G., 2021. Forecasting: principles and practice, 3rd ed. OTexts: Melbourne, Australia.

[15] Brockwell, P.J., Davis, R.A., 1991. Time series: theory and methods, 2nd ed. Springer New York, NY.

[16] Smejkalová, V., Šomplák, R., Nevrlý, V., Burcin, B., Kučera, T., 2020. Trend forecasting for waste generation with structural break. Journal of Cleaner Production 266.

[17] Harvey, A.C., Shephard, N. Structural time series models, in: Handbook of Statistics, vol. 11, pp. 261–302.

[18] Livera, A.M. de, Hyndman, R.J., Snyder, R.D., 2011. Forecasting time series with complex seasonal patterns using exponential smoothing. Journal of the American Statistical Association 106 (496), 1513–1527.

[19] Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32.

[20] Grinsztajn, L., Oyallon, E., Varoquaux, G. Why do tree-based models still outperform deep learning on typical tabular data?, in: NIPS'22: Proceedings of the 36th International Conference on Neural Information Processing Systems, pp. 507–520.

[21] Huang, H., Pouls, M., Meyer, A., Pauly, M. Travel time prediction using tree-based ensembles, in: Lalla-Ruiz, E., Mes, M., Voß, S. (eds) Computational Logistics, vol. 12433, pp. 412–427.

[22] Jayaraman, V., Parthasarathy, S., Lakshminarayanan, A.R., Singh, H.K. Predicting the quantity of municipal solid waste using XGBoost model, in: Third International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 148–152.

[23] Cubillos, M., 2020. Multi-site household waste generation forecasting using a deep learning approach. Waste Management 115, 8–14.

[24] R Core Team, 2025. R: a language and environment for statistical computing, Vienna, Austria. https://www.R-project.org/. Accessed 30 March 2026.

[25] Fahrmeir, L., Kneib, T., Lang, S., Marx B.D., 2021. Regression: models, methods and applications, 2nd ed. Springer Berlin, Heidelberg.

[26] Wright, M.N., Ziegler, A., 2017. ranger: A fast implementation of Random Forests for high dimensional data in C++ and R. Journal of Statistical Software 77 (1), 1–17.

[27] xgboost 3.1.2.1 documentation, 2025. XGBoost parameters. https://xgboost.readthedocs.io/en/latest/parameter.html#. Accessed 30 March 2026.

Biography

Leia Betting (*1999) studied Statistics at TU Dortmund University, completing her Bachelor's and Master's (2019–2025). Her Bachelor thesis received the GMDS Förderpreis 2023. Her Master thesis focused on forecasting in the returnable bottle sector. Since October 2025, she is a research associate and Ph.D. student at the Statistics Department of TU Dortmund University within the TRR 391 "Spatio-temporal Statistics for the Transition of Energy and Transport".

Jakob Becker (*1999) studied Mathematics and Computer Science at the Heinrich-Heine University Düsseldorf in his Bachelor's and Statistics & Data Science at the Ludwig-Maximilians University Munich in his Master's. In his current Ph.D. studies, he holds a scholarship at the Graduate School of Logistics at TU Dortmund University, where analyzes recycling processes with statistics and machine learning. ORCID: 0009-0004-0694-590X

Philipp Kempkes (*1976) received his diploma degree in physics from Ruhr-University Bochum in 2005 and completed his Ph.D. in experimental plasma physics in 2008. After working as a Post-Doc at Ruhr-University Bochum and Princeton Plasma Physics Lab, he was an associate professor at the Max-Planck-Institute for plasma physics at the University of Greifswald. Since 2014, he works as team lead for digitalization and R&D at REMONDIS.

Konrad Philipp Mielke (*1992) has studied physics at TU Dortmund University (2012-2017) and completed his Ph.D. in data science at Radboud University Nijmegen in 2024. Since 2022, he is working at REMONDIS on the topics of digitization and research & development.

Markus Pauly (*1980) received his diploma degree in mathematics from the Heinrich-Heine University Düsseldorf (2005), where he also completed his Ph.D. (2008) and his habilitation (2013). From 2014 to 2019 he was full professor for Statistics at Ulm University. Since 2019 he is full professor at the Statistics Department at TU Dortmund University, investigating a wide range of statistical and machine learning topics. ORCID: 0000-0002-0976-7190

8th Conference on Production Systems and Logistics

Digital Transparency In The Loop Between Shopfloor And Logistics Planning

Julian Hartwich-Fischer[1], Katharina Paschert[2], Anja Munzke[1], Mario Luber[1], Florian Karg[1], Meltem Akkaya[1]

[1]*Fraunhofer Institut for Casting, Composite and Processing Technology IGCV, Augsburg, Germany*
[2]*Ingenics AG, Ulm, Germany*

Abstract

Automotive companies are confronted with significant challenges, including rising customer expectations regarding sustainability, intensified global competition, and the transition toward electromobility. Stable processes, high efficiency, and scalability can only be achieved if relevant target values such as energy consumption, throughput time, and costs are defined in the planning phase and continuously optimized during the operational phase. Reaching these objectives necessitates the establishment of digital transparency, which enables the utilization of data from physical resources and information flows for planning purposes. This approach facilitates the integration of real data from the shopfloor into the planning of new production systems and the optimization of existing processes, thereby ensuring a data-driven decision-making process. However, the successful implementation of this approach depends on the identification of the necessary data, the development of suitable data evaluation methods, and the incorporation of this data into the planning process in an appropriate form. Consequently, planning projects often suffer from a lack of end-to-end digital transparency, in which the data necessary for planning or optimization is also available to achieve optimal planning results. The presented paper outlines the use of a digital resource twin, which integrates logistics resources with planning, thereby facilitating digital transparency between planning and operations. Additionally, it provides an exposition of three scenarios in which the system can be employed to optimize the economic and ecological performance of logistic processes within the automotive industry. The system developed is a purpose-driven innovation that integrates existing technologies to enhance efficiency in the field of intralogistics. This paradigm shift introduces an entirely novel framework for the planning of future production systems. Moreover, it facilitates the effective integration of planning results into the control systems that orchestrate operations on the shopfloor.

Keywords

Digital Transparency; Intralogistics; Factory Planning; Digital Twin; Data Integration; Process Optimization; Automotive Industry

1. Introduction

The increasing complexity, ever-shorter innovation cycles, growing variety of products, and rising individualization of passenger vehicles have significant effects on the production and intralogistics infrastructures of the automotive industry and its suppliers [1]. Moreover, the sector is confronted with distinctive challenges, including highly volatile markets and substantial cost pressures in the context of globalisation [2,3]. These developments are particularly evident in the sales figures for battery electric vehicles [4]. Especially during the ongoing transformation process in the context of the electrification of

DOI: https://doi.org/10.15488/21017
ISSN: 2701-6277

drivetrain technologies over the coming decades, a multitude of vehicle variants will need to be produced simultaneously, further increasing the complexity of production environments [5]. However, this increase in complexity has implications that extend beyond the complexity of production environments; it also affects intralogistics processes. Efficient material supply is a critical factor in maintaining continuous production and minimizing operational disruptions in manufacturing systems [6,7]. Within automotive manufacturing, the timely replenishment of production lines constitutes a pivotal element of the production system. Here, delays and supply bottlenecks due to inadequate route planning can lead to frequent production stoppages, increased operational costs, and reduced competitiveness [8]. This underscores the significance of intralogistics planning as a critical component in the planning of efficient production systems. However, the quality of intralogistics planning results is currently being adversely affected by a lack of data for process design. The lack of data represents a significant challenge across the domain of factory planning and can therefore act as an obstacle to intralogistics planning [9]. Moreover, most companies persist in prioritizing data collection, while the targeted use of data remains infrequent.

Concurrently, the utilization of shopfloor data can serve as a pivotal instrument for enhancing the efficiency of intralogistics processes from the planning stage onwards [10]. The present work builds on the development by Luber et al. and introduces a concept for the utilization of digital resource twins (DRT) for intralogistics planning, thereby facilitating digital transparency between planning and operations while enhancing the planning quality [11]. In this regard, Chapter 2 initiates with an exposition of digital intralogistics planning as a distinct domain of responsibility within the overarching framework of factory planning. The chapter also presents the concept of a DRT. The third chapter then provides a comprehensive explanation of the concept, addressing its individual components in detail. The subsequent fourth chapter provides an exposition of three specific use cases in which the developed concept can be applied in the domain of intralogistics planning. Finally, Chapter 5 provides a brief summary and conclusion.

2. Fundamentals and research gap

2.1 Digital intralogistics planning

Intralogistics describes the complex cooperation of different logistics functions and includes all internal flows from organizing, controlling, executing, and optimizing all activities within a unit like a factory or distribution centre, with the goal to provide material in the demanded quality. In the context of this paper dealing with the automotive industry, intralogistics is seen as a subdiscipline of a factory, both in planning and operational matters [12,13]. Intralogistics operations include material flow and information flow processes, providing the right material at the right time in the right quality and the right place. The material flow is executed by material handling equipment used for storage, movement, assortment, and management, that builds a network of logistics resources collaborating on the shopfloor. These resources are steered and controlled by logistics IT systems, e.g., transport management systems (TMS), and are embedded in the local infrastructure [14,15]. Intralogistics planning is part of factory planning, the overall process of which is often divided into several phases, e.g., described in the VDI guideline 5200, that provides a basis for methodical quality [13]. Adapted into a more practical and established approach, intralogistics planning is often divided into analysis, concept, and realization phases, extended by equipment tendering, operations accompaniment, or change management [16,17]. One core task is the material flow planning, that combines efficient processes with best-fit handling equipment. Depending on the requirements and general conditions, making the right decisions in layout allocation, transport routes, and equipment procurement affects the quality of a production system. Requirements could be e.g., the demanded level of automation, sustainability, or flexibility targets, or the integration in a grown brownfield infrastructure [14,12]. The planning process itself is well established but faces increasing challenges in the context of digital transformation These challenges include the imperative for real-time decision-making to achieve process excellence and the absence of collaboration between factory functions [18,15].

At present, a substantial number of software solutions are available for use in the various tasks associated with factory planning. According to the study published by Bermpohl et al., computer-aided design (CAD) software, enterprise resource planning systems, and simulation tools have been identified as being among the most frequently used tools in factory planning [9]. In the domain of intralogistics planning, material flow simulation can facilitate the optimization of planning outcomes and guarantee the requisite performance of the intended logistics system [19]. In the context of this paper, material flow simulation is the "representation of a system with its dynamic processes in an experimental model to reach findings which are transferable to reality" [20]. In a logistics planning process, options for layouts, processes, or handling equipment can be modelled in a simulation tool. Through simulation experiments, results and variables are collected and evaluated, based on statistical methods. This approach facilitates the comparison of various scenarios and thereby contributes to the optimization of the planning process and the selection of the optimal variant. In addition to its role in planning, simulation can be utilized during operations for the purpose of conducting parameter studies and assessing the potential for optimization [20]. Additionally, simulation is not only suitable for evaluating material flow processes, but also for e.g. coupled production and energy systems [21]. A core for successfully leveraging simulations is the (partially) automated processing of input data in terms of digital information consistency [22]. Simulation-based logistics planning is already known [19,23]. But without real-time data, high-performance interfaces or deep level knowledge of the complex shopfloor network, planning processes are still inefficient and prone to errors, not yet Industry-4.0-ready [24,25].

2.2 Digital resource twin

A digital twin in production is a "fit for purpose digital representation of an observable manufacturing element with synchronization between the element and its digital representation" [26]. A digital twin's purpose is to support decision making and provide relevant descriptive, predictive, or prescriptive models, which reflect characteristics and behavioral properties of the real object [27]. These can be processing times, remaining usage durations, or availability ratios, dependent on specific production conditions in the case of a production resource. The synchronization to the real object is typically achieved through a continuous data connection, which provides the object's condition data to the digital twin. Based on these actual measured data, the decision supporting models are updated or adapted for an up-to-date and evidence-based virtual representation [28]. Digital twins are often enhanced by simulation capabilities and a bidirectional data connection that allows for an automated execution of the induced decision [29]. Digital twins represent entire systems, e.g., an assembly line or a material supply system, or subordinate resources, such as an assembly station or a transport vehicle. This paper focuses on the use of DRTs.

Luber et al. propose the concept of a ready-to-use, data analytics-based DRT, which uses the data acquired on the shopfloor to create predictive models for planning applications in manufacturing [11]. Luber et al. develop and demonstrate a corresponding software architecture. These novel concepts and developments constitute the foundation of the DRT, which is applied in the context of logistics planning for the automotive sector in this paper [30]. A key element is the ready-to-use aspect: pre-configured, generally applicable, and interoperable components ensure a low implementation effort. This aims at overcoming the main barrier of deploying Digital twins. Figure 1 outlines the software architecture of the DRT. The configurable shopfloor client establishes connections to data sources, such as an OPC UA server or Ethernet fieldbus protocols. A resource-specific algorithm transforms the connected condition data into the defined structure of the shopfloor records service of the DRT. The data is then transferred to the deployed Digital twin server. A pre-configured Data Analytics algorithm uses the prepared data records to create the predictive models, which are offered to planning applications by dedicated services. Metadata required to call the services, as well as the foundational shopfloor records, are stored in the Asset Administration Shell (AAS) of the resource. The AAS ensures interoperability of the DRT by defined connectors based on REST/HTTP and data schemes. Prerequisite for the application of a DRT is the accurate measurement of shop floor data. While individual

faulty shop floor data, which is outside a valid value range, can be dropped during data transformation, an initial assessment of the general validity of the acquired data has to be carried out.

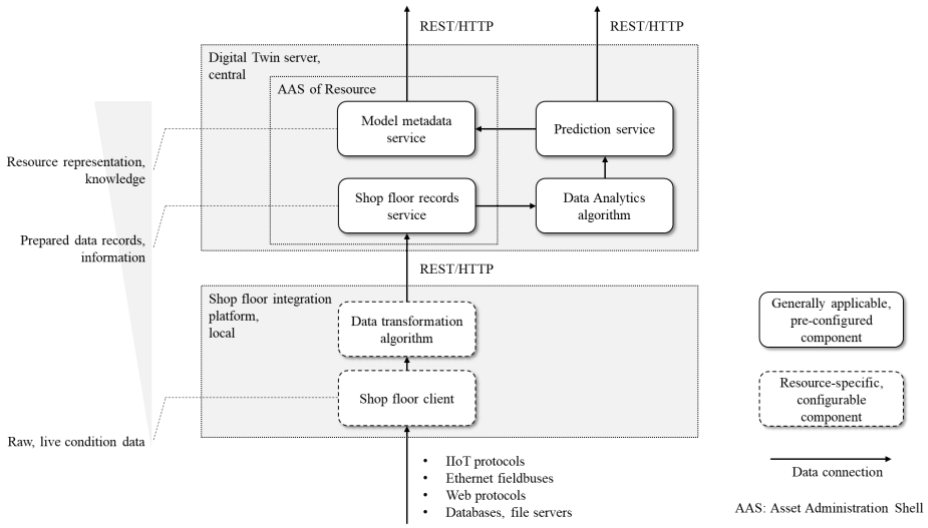


Figure 1: Software architecture of the DRT [30]

2.3 Research gap

Currently, the potential application scenarios for the DRT presented by Luber et al. have yet to be explored in the domain of intralogistics planning [11]. As a result, intralogistics planning and material flow simulation remain predominantly reliant on manufacturer specifications. As previously mentioned in the preceding chapters, in the context of planning projects, insufficient databases pose a significant challenge, with the quality of simulation results being largely contingent upon the parameters that are established and the input data that is utilized [24,9]. The concept developed aims to counteract precisely this problem. The DRT is integrated with planning applications to enable the use of fit-for-purpose, processed shopfloor data that supports planning tasks. The interconnection between resources on the shopfloor and material flow simulation as a tool in intralogistics planning establishes a digital flow of information from operations to planning. The subsequent chapter presents a novel concept for leveraging DRT, thereby enabling the automated configuration of material flow simulation parameters derived from shopfloor data.

3. Concept for using digital transparency in intralogistics planning

3.1 Concept structure

The present paper proposes a concept for the optimization of intralogistics planning through the utilization of shopfloor data, with the objective of enhancing the quality of planning results. The concept has been demonstrated to enhance the digital transparency of logistics processes and to provide a methodology for the complete digitalization of information flows from the shopfloor to factory planning. As illustrated in Figure 2, the overarching concept is comprised of two distinct levels, each of which fulfills a specific function and is interconnected through the implementation of a DRT. In essence, the shopfloor level constitutes the data source, while the factory planning level signifies the target area in which the measured data is to be utilized for planning tasks. The DRT developed by Luber et al. corresponds to the structure explained in Chapter 2.2 and facilitates the integration of these two levels [11]. Furthermore, the DRT is utilized in this concept to transform shopfloor data and render it usable for planning purposes, thereby optimizing planning quality. The subsequent subchapters will examine the shopfloor level and the factory planning level, as well as their respective functional units, in detail.

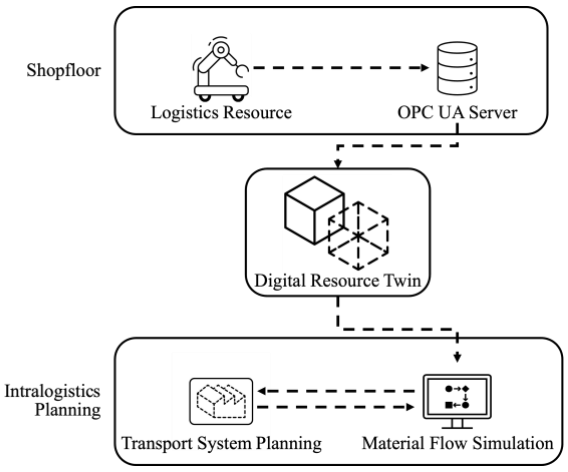


Figure 2: Concept for the establishment of digital transparency between shopfloor and factory planning

3.2 Concept components

3.2.1 Shopfloor

The shopfloor level comprises all operational elements of intralogistics that execute and monitor the physical flow of materials in the plant, as well as the associated information systems. These include conveyor systems, tugger trains, mobile robots, forklifts, supermarkets/picking zones, and load carriers, as well as identification technology. From a systems perspective, manufacturing execution systems, warehouse management systems, material flow systems, and control stations are of relevance. The shopfloor level is characterized by the continuous generation and consumption of data streams, including transport orders, inventory and demand information, plant and vehicle statuses, positions and movements, energy and battery data, safety statuses, quality and process events. A subset of the data is event-based and high-frequency, while another subset consists of master data and shift schedules. This data must be synchronized in terms of time, harmonized semantically, and quality-assured (completeness, plausibility, time stamp) before it can be made available. Restrictions on the shopfloor delineate the permissible solution space for material flows and reflect real-world conditions. Concurrently, disruptions and variabilities (e.g., blockages, rush orders, delays) function as exogenous events that must be evident in the data image to facilitate the establishment of robust and realistic planning assumptions. In the context of material flow simulation, the shopfloor level thus functions as a reference and control variable, thereby facilitating the provision of real data. Furthermore, approved planning artifacts, including route and scheduling rules, fleet sizes, loading, and buffer management, can be integrated into operational execution. When establishing a logistics resource connection, it is imperative to distinguish between subsystems. Each subsystem provides relevant operating and status data to a server, which then consolidates this data for the overall system. As illustrated in Figure 3, a potential logistics resource could be a mobile manipulator comprising three distinct subsystems: a mobile platform, a manipulator, and a gripper.

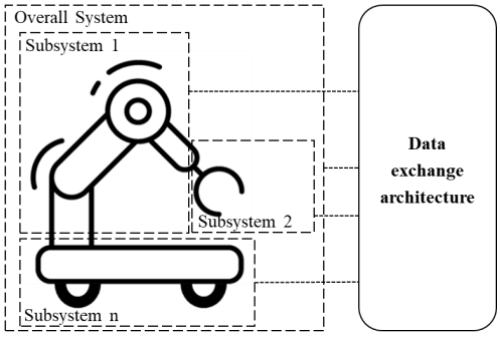


Figure 3: Subsystems at shopfloor level

A comprehensive dataset is collected, including location, orientation, velocity, axis/position data, energy consumption, and a unified operating status (Not Available, Out of Service, Not Executing, Executing). This calls for a networked control system that conforms to industry standards, such as a programmable logic controller (PLC) linked to appropriate sensors. An OPC UA server can function as a standardized data hub, consolidating operating/status data into a semantically harmonized model and enabling control interventions. As previously outlined in section 2.3, the OPC UA server can be utilized to establish a connection for data transfer with the DRT. The data is transferred to the node set of the DRT, which functions as a continuous interface between the shopfloor, planning, and simulation. There, the collected data undergoes a transformation process, resulting in values that can be seamlessly processed at the factory planning level.

3.2.2 Intralogistics planning

The objective of the presented concept is to enhance the quality of planning results by leveraging the recorded data of specific logistics resources. In the context of this concept, material flow simulation constitutes an integral component of intralogistics planning. In the simulation model, the real logistics process is mapped using the same resource whose operating data is collected, aggregated, and transformed during ongoing operations. However, it is important to note that this does not necessarily imply that the process must be identical; it is sufficient to utilize an identical logistics resource. A set of parameterizable characteristics (e.g., speed and technical availability) is defined for each resource. To establish a fully digital process for resource parameterization, a connection was established between the material flow simulation and the DRT. Accordingly, the parameters are transferred automatically from the administration shell using a method implemented in the simulation model. These parameters correspond to the values measured on the shopfloor. The implemented method uses dynamically triggered HTTP requests based on the operating status of the resource. For each detected state, a POST request containing the relevant parameters in JSON format is sent to a predetermined endpoint of the administration shell. The AAS then returns with state-specific characteristic values, which are then processed within the simulation model. This state-dependent, HTTP-based data exchange enables the simulation to operate continuously with accurate, up-to-date values. This quality-assured data is the foundation upon which simulation results are derived, reflecting the behavior of the real logistics resource with greater precision. In intralogistics planning, the resulting low-distortion results can then be used for the design and optimization of logistics processes.

4. Concept validation

To validate the concept, three use cases were defined, each targeting specific application scenarios and using data from a shopfloor resource. A common intralogistics process is exhibited by all three. In this process, an autonomous mobile robot (AMR) retrieves small load carriers (SLC) from storage and delivers them to multiple locations. The process flow is as follows: the AMR begins in a designated waiting position, subsequently moves to the pickup location, loads SLCs of varying weights, follows a predetermined route to the destinations, unloads at each location, and returns to its waiting position. The use cases were selected to investigate the potential of data-driven approaches in different operational and planning contexts, while taking supply security requirements into account. These requirements are a fundamental prerequisite for all three scenarios. The three use cases were derived from the common objectives within intralogistics planning [31,20]. The first use case targets reducing energy consumption by recording AMR process and status data, enabling energy-optimized planning. The second use case examines how the flexibility of a logistics system's volume can be increased by integrating shopfloor data into logistics planning. The third use case focuses on analyzing and optimizing operational key figures. Available data is used to determine relevant performance indicators more precisely, identify deviations at an early stage, and support data-driven improvement measures. Specifically, the aim is to record, analyze, and use disruptions in operational processes as a basis for optimization. To validate the use cases, the components of the concept were implemented as demonstrators to verify functionality and effectiveness under realistic conditions. It is important to note that

Use Cases 1 and 2 are externally triggered. Use Case 1 instigates a deliberate optimization, while Use Case 2 leads to a deliberate rescheduling. Use Case 3 is internally triggered when continuous KPI monitoring detects insufficient buffer supply and sub-threshold fill levels. The subsequent sections will present each use case with its objective and hypothesis, followed by an examination of the initial situation, execution, and results.

4.1 Use Case 1: Analysis of energy reduction through planning measures

The first use case aims to demonstrate the potential for reducing energy consumption through planning measures in material flow systems. The starting point of the investigation is the following hypothesis:

The use of real operational data in terms of load an energy consumption results in a different assembly of transport routes compared to planning based on manufacturer-provided data.

To validate the hypothesis, the initial situation was reproduced in a demonstrator. The demonstrator consisted of an AMR that ran a three-stop route. The predefined route was determined to be the shortest by distance. The AMR picks up SLCs of varying weights and continuously transmits measured, state-dependent energy consumption and current load to the DRT. A material-flow simulation is employed to replicate this setup and ingests the demonstrator's real data via an interface, as delineated in Chapter 3. Within the simulation environment, alternative route options as well as different load combinations were subsequently investigated. The evaluation of the simulation results shows that, for the same transport task, energy demand can increase due to a different route selection, even if all destination stations are still served. For comparison, the same transport task was additionally simulated using the energy data provided by the manufacturer. Here as well, predicted deviations in energy consumption could be demonstrated, underlining the relevance of using real, measured consumption data. When these findings are transferred to a complex transport network, it becomes evident that both adapted route planning and a targeted design of the load represent essential levers for reducing energy demand.

4.2 Use Case 2: Increasing volume flexibility through the integration of shopfloor data into logistics planning

The second use case investigates the potential to increase the volume flexibility of a logistics system by integrating shopfloor data into the planning, and particularly the dimensioning of ARM systems. The study is based on the following hypothesis:

The use of real operational data results in a different demand for AMRs compared to planning based on manufacturer-provided data.

To enhance volume flexibility, an increase in the transport capacity of the system is required. In this context, state-dependent AMR speeds are a critical factor in determining transport performance and are therefore explicitly considered in the planning process. The recorded real speed profiles are integrated into the material flow simulation to support planners in designing the transport system for increased throughput requirements. Based on this simulation, the required number of AMRs is determined, while additionally defining a target corridor for vehicle utilization. This approach ensures two things: first, that sufficient transport capacity is available to cope with volume fluctuations, and second, that an economically reasonable utilization of the AMRs is maintained. Under the investigated configuration, the increased transport demand leads to a higher number of required AMRs. As a reference scenario, the transport system is additionally designed using the average speeds specified by the manufacturer. A comparison of the two planning approaches reveals that assuming constant, idealized speeds can result in an underestimation of the required fleet size. In real operation, the speeds vary with vehicle dynamics (e.g., cornering) and disturbances, replicated in the laboratory setting (pallet obstacles, pedestrian traffic). By incorporating these influences into the simulation, the real system behavior can be represented much more accurately, enabling a more robust design of the transport system with respect to vehicle dimensioning.

4.3 Use Case 3: Using real operational data to optimize logistic processes

The third use case aims to improve operational performance through the targeted optimization of logistics processes. The primary focus is on identifying and reducing excessive lead times in the supply of buffers at the production facilities. To this end, the following hypothesis is proposed:

The use of real operational data enables the systematic identification of operational disturbances and their integration into a simulation model as planning parameters, enabling deviations from predefined operational target values to be detected at an early stage.

During operation, operational disturbances are deliberately introduced—following the approach of Use Case 2—to represent realistic shopfloor conditions. In contrast to the uniform adjustment of the average travel speed applied in Use Case 2, this use case focuses on the identification, classification, and model-based representation of different types of disturbances. Based on operational data collected on the shopfloor, specific disturbance profiles are derived, which can be characterized by their frequency of occurrence and duration. If these disturbances cannot be eliminated through immediate operational measures, they are incorporated into the simulation as fixed planning parameters. Using the extended simulation model, it is subsequently analyzed how many additional AMRs are required to reliably fulfill the transport tasks despite the identified disturbances and to ensure a continuous supply of the buffers at the production facilities. At the same time, the utilization of the AMRs is monitored. In cases of insufficient utilization, additional transport tasks are assigned to the AMRs while explicitly accounting for operationally induced disturbances, thereby ensuring a realistic and robust deployment strategy.

5. Summary and conclusion

Intralogistics constitutes a pivotal element within the framework of any production system, exerting a substantial influence on its performance. However, the complex environment and frequently inadequate data basis complicate the achievement of optimal intralogistics process planning. Nevertheless, innovative approaches to data utilization and digital planning instruments have the potential to facilitate the resolution of this challenge. In this context, the present paper introduces a concept that utilizes a DRT to facilitate the automated transfer of shopfloor data to a material flow simulation and, consequently, to the planning level. This establishes a digital infrastructure that fosters transparency between the operational floor and the planning level. In addition, three use cases were employed to illustrate the application of the presented concept in enhancing planning quality and, consequently, optimizing intralogistics processes in a targeted manner. The transferability of the concept is contingent upon the definition of the data to be collected from the logistics resource. The determination of these parameters must be derived from a specific use case to optimize a specific target variable within the context of intralogistics planning. In the next step, the concept shown must be verified in the defined use cases. To this end, the components of the overall system have been installed in demonstrator form and implemented in a laboratory environment. In addition, the use case delineated in Chapter 4.1 has undergone testing and validation, while the requisite data for the remaining two use cases has been defined.

6. License agreement

In order for papers to be published via the repository of the Leibniz Information Centre for Science and Technology and University Library (TIB), the authors must transfer a right of use for their paper to the operators of the repository (Leibniz University Hanover and TIB) - for online provision in the repository. By submitting your paper, you agree to accept the license agreement or authorize the organizers of the conference to accept it for you. For more information, visit the TIB repository at **https://www.repo.uni-hannover.de**.

Acknowledgements

Special thanks go to the German Federal Ministry of Economic Affairs and Energy (BMWE) for funding the research project “DiMoLo” | Digital transparency as an enabler for the design of economically sustainable assembly and logistics systems for the future

References

[1] Winkelhake, U., 2025. The Digital Transformation of the Automotive Industry. Springer Berlin Heidelberg, Berlin, Heidelberg.

[2] Lempp, M., Siegfried, P., 2022. Automotive Disruption and the Urban Mobility Revolution. Springer International Publishing, Cham.

[3] Siegl, L., Bornemann, T., 2025. Assessment of shop floor environment dynamics in production plants by developing an innovative methodology for an appropriate implementation of mobile transport robots. Prod. Eng. Res. Devel. 19 (6), 1231–1246.

[4] Puls, T., 2024. Die Automobilindustrie im Jahr 2024: Globale Trends stellen das erfolgreiche Geschäftsmodell der deutschen Autoindustrie vor Herausforderungen. https://www.iwkoeln.de/fileadmin/user_upload/Studien/Report/PDF/2024/IW-Report_2024-Automobilindustrie.pdf.

[5] Colovato, M.R., Baldo, C.R., 2023. Industry 4.0 Support to Automotive Electrification: Current Challenges and Opportunities, in: 2023 15th IEEE International Conference on Industry Applications (INDUSCON). 2023 15th IEEE International Conference on Industry Applications (INDUSCON), São Bernardo do Campo, Brazil. 22.11.2023 - 24.11.2023. IEEE, pp. 1213–1218.

[6] Glova, J., Sabol, T., Vajda, V., 2014. Business Models for the Internet of Things Environment. Procedia Economics and Finance 15, 1122–1129.

[7] Ojstersek, R., Javernik, A., Buchmeister, B., 2021. The impact of the collaborative workplace on the production system capacity: Simulation modelling vs. real-world application approach. Adv produc engineer manag 16 (4), 431–442.

[8] Pekarčíková, M., Kliment, M., Kronová, J., Trebuňa, P., Hovana, A., 2025. Simulation-Based Optimization of Material Supply in Automotive Production Using RTLS Data. Applied Sciences 15 (16), 9102.

[9] Bermpohl, F., Schäfer, S.F., Neumann, O., Reihlen, E., Dickopf, T., Gebel, T., Neuhäuser, T., Daub, R., 2025. Industrial study on holistic digital factory models. Prod. Eng. Res. Devel.

[10] Olewe, S., Finke, M., Belke, J., Dyck, F., Kürpick, C., 2023. Use Case Catalog and Assessment for AI Applications in Intralogistics of Manufacturing Companies. Procedia CIRP 118, 74–79.

[11] Luber, M., Wagner, S., Wegmann, M., Schilp, J., 2025. Concept of a plug-and-play, Machine Learning Digital Twin of the production resource for Detailed Capacity Planning and Scheduling. Procedia Computer Science 253, 1266–1275.

[12] Fottner, J., Clauer, D., Hormes, F., Freitag, M., Beinke, T., Overmeyer, L., Gottwald, S.N., Elbert, R., Sarnow, T., Schmidt, T., Reith, K.B., Zadek, H., Thomas, F., 2021. Autonomous systems in intralogistics-state of the art and future research challenges. Logistics Research 14 (1).

[13] Munzke, A., Neuhäuser, T., Hohmann, A., Daub, R., 2022. Efficient Intralogistics Planning Based on an Innovative Intralogistics Tool using the Example of a Flexible Battery Cell Factory, in: , Proceedings of the Conference on Production Systems and Logistics: CPSL 2022. Hannover : publish-Ing, pp. 1–10.

[14] Božić, M., Đurđević, D., Dabić-Miletić, S., 2024. Guidelines for technology selection in intralogistics: A scientific approach. Vojnotehnički glasnik 72 (4), 1948–1976.

[15] Perman, Dovletgeldi, Begench, Bayramgeldi, Tahyr, 2025. Digital Transformation in Logistics: Key Drivers, Challenges, and Sustainable Solutions for the Modern Supply Chain, 9 pp. Accessed 8 December 2025.

[16] Intralogistik Planung - Fraunhofer IML. https://www.iml.fraunhofer.de/de/abteilungen/b1/intralogistik-und--it-planung/dienstleistungen0/intralogistikplanung.html. Accessed 8 December 2025.

[17] Fottner, J., Galka, S., Habenicht, S., Klenk, E., Meinhardt, I., Schmidt, T., 2022. Planung von innerbetrieblichen Transportsystemen. Springer Berlin Heidelberg, Berlin, Heidelberg.

[18] Butt, A.S., Alghababsheh, M., Awan, U., 2025. Barriers to the adoption of the Metaverse in digital supply chains. JEIM.

[19] Golova, J., Mahmood, K., Raamets, T., 2021. Simulation based Performance Analysis of Production Intralogistics. IOP Conf. Ser.: Mater. Sci. Eng. 1140 (1), 12026.

[20] Verein deutscher Ingenieure, 05 2018. VDI 3633 Blatt 1 : Simulation von Logistik-, Materialfluss- und Produktionssystemen - Grundlagen. VDI-Verlag, Düsseldorf.

[21] Sokoll, A., Buning, B., Schäfer, S.F., Dickel, K., Herrmann, C., 2025. Discrete-Event Simulation for improving coupled production and energy systems, in: Simulation in Produktion und Logistik 2025 21. ASIM-Fachtagung. 21. ASIM-Fachtagung Simulation in Produktion und Logistik. 24. bis 26. September 2025. Technische Universität Dresden.

[22] Schäfer, F.S., Buning, B., Krüger, T., Bermpohl, F., Münnich, M., 2025. Verbindung von BIM und Simulation/Linking BIM and simulation - Motivation and existing challenges in factory planning. wt 115 (06), 517–525.

[23] Mahmood, K., Karjust, K., Raamets, T., 2021. Production Intralogistics Automation Based on 3D Simulation Analysis. Journal of Machine Engineering, 101–115.

[24] Albrecht, T., Baier, M.-S., Gimpel, H., Meierhöfer, S., Röglinger, M., Schlüchtermann, J., Will, L., 2024. Leveraging Digital Technologies in Logistics 4.0: Insights on Affordances from Intralogistics Processes. Inf Syst Front 26 (2), 755–774.

[25] Siva Ramkumar, M., Arun, M., Gopan, G., Giri, J., Al-Qawasmi, K., 2025 - 2025. Revolutionizing Global Logistics: Integrating Advanced Supply Chain Management Techniques in the Digital Era, in: 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI). 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI), Erode, India. 1/20/2025 - 1/22/2025. IEEE, pp. 79–85.

[26] ISO 23247. Automation systems and integration - Digital twin framework for manufacturing.

[27] Ivanov, D., 2024. Conceptualisation of a 7-element digital twin framework in supply chain and operations management. International Journal of Production Research 62 (6), 2220–2232.

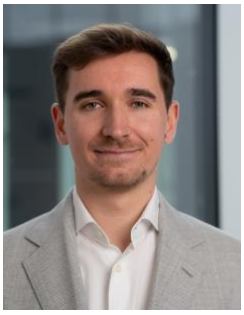
[28] Wright, L., Davidson, S., 2020. How to tell the difference between a model and a digital twin. Adv. Model. and Simul. in Eng. Sci. 7 (1), A6.

[29] Negri, E., Fumagalli, L., Macchi, M., 2017. A Review of the Roles of Digital Twin in CPS-based Production Systems. Procedia Manufacturing 11, 939–948.

[30] Luber, M., Wittmann, L., Schilp, J., 2026. Software infrastructure for ready-to-use, data analytics-based Digital Twins utilising the Asset Administration Shell. Procedia Computer Science.

[31] Gudehus, T., 2010. Logistik: Grundlagen-Strategien-Anwendungen. Springer Berlin Heidelberg, Berlin, Heidelberg.

Biography



Julian Hartwich-Fischer (*1999) is a research associate at the Fraunhofer Institute for Casting, Composite and Processing Technology IGCV and works since 2024 in the group of collaborative factory planning. He studied industrial engineering with a focus on mechanical engineering at the Technical University of Ilmenau and Friedrich-Alexander University Erlangen-Nuremberg.

Katharina Paschert (*1992) is Consultant for Logistics and Supply Chain Management at Ingenics AG since 2018. Since 2024 she has the role of Team Manager for Logistics Technology. She studied Economic Logistics at the University of Applied Sciences in Dortmund (B.Sc.)

Anja Munzke (*1993) studied systems engineering and management with a specialization in logistics at the University of Applied Sciences THU in Ulm. 2019 she started working at the Fraunhofer Institute for Casting, Composite and Processing Technology (IGCV). Since 2024 Anja Munzke is head of department for sustainable factory planning and operation at the Fraunhofer IGCV.

Mario Luber (*1989) studied Mechanical Engineering and Management at the Technical University of Munich. After completing his master's degree, he worked as an IT application consultant before joining the Digital Production and Artificial Intelligence department of the Fraunhofer Institute for Casting, Composite and Processing Technology (IGCV).

Florian Karg (*1997) studied mechatronics with a focus on robotics at the University of Applied Sciences in Augsburg and the University of Ulster. In 2021, he started working at the Fraunhofer Institute for Casting, Composite and Processing Technology (IGCV). He specializes in long-term autonomous robotics, especially mobile manipulators.

Meltem Akkaya (*1998) is a research associate at the Fraunhofer Institute for Casting, Composite and Processing Technology IGCV and works since 2024 in the group of collaborative factory planning. She studied industrial engineering with a focus on management and sustainabilityat the University of Augsburg.

8th Conference on Production Systems and Logistics

Methodology For Dynamic Dimensioning Of Re-Assembly Factories

Günther Schuh[1,2], Esben Schukat[1], Nils Lehde genannt Kettler[1]
[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany*
[2]*Institute/ Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

The transformation toward a circular economy requires the industrial scaling of re-assembly processes. However, re-assembly factories face significant uncertainties regarding product condition, variant diversity, and volatile demand, which push existing factory planning approaches to their limits. This dissertation addresses the central challenge of dynamic dimensioning of re-assembly factories – the continuous adaptation of capacities and production structures to changing boundary conditions. In contrast to conventional concepts of flexibility and changeability, which exhibit either insufficient scopes of change or inadequate adaptation speeds, this work develops a methodology for proactive, data-driven dimensioning. The overall solution consists of three essential modules: First, the digital product passport that contains various data classes with different levels of completeness, detail, and quality. Second, a model of cause-effect relationships that links product data to capacity planning requirements. Third, capacity planning methods ranging from rough estimation to second-precise operational details. The core of the solution is the systematic utilization of digital product passports to derive ReX strategies that serve as a link between product requirements (demand) and production system configuration (supply). The methodology encompasses the development of a descriptive model of strategic premises, an explanatory model of cause-effect relationships, and a decision model for adaptive capacity planning and production segmentation. This work contributes to enabling industrial re-assembly factories to operate in an economically scalable manner, thereby significantly supporting the transformation toward a circular economy.

Keywords

Factory Planning; Capacity Planning; Digital Product File; Re-Assembly; Circular Economy

1. Introduction

Industrial production is currently characterized by linear value creation [1]. This principle leads to the wealth of industrial society and the consumption of natural resources increasing in parallel with one another [2]. Linear value creation does not have an economic or ecological future due to planetary boundaries [3]. The circular economy offers the potential to uncouple resource consumption and economic growth [4]. The simplified goal of the circular economy is to absolutely reduce resource and energy consumption [5]. The circular economy provides various ReX strategies to achieve this goal. These ReX strategies either involve making sensible use of processed raw materials, extending the life cycle of products, or manufacturing and using products in a smarter way.[6] One of these ReX strategies is re-assembly. This ReX strategy is a further development of remanufacturing. The re-assembly strategy involves the industrial upgrading of products by replacing, reprocessing, or upgrading components and providing them to customers as new products.[7] The categorization and differentiation of re-assembly from other ReX strategies can be described in terms of level and result of value preservation. The categorization is shown in Figure 1.

DOI: https://doi.org/10.15488/21018
ISSN: 2701-6277

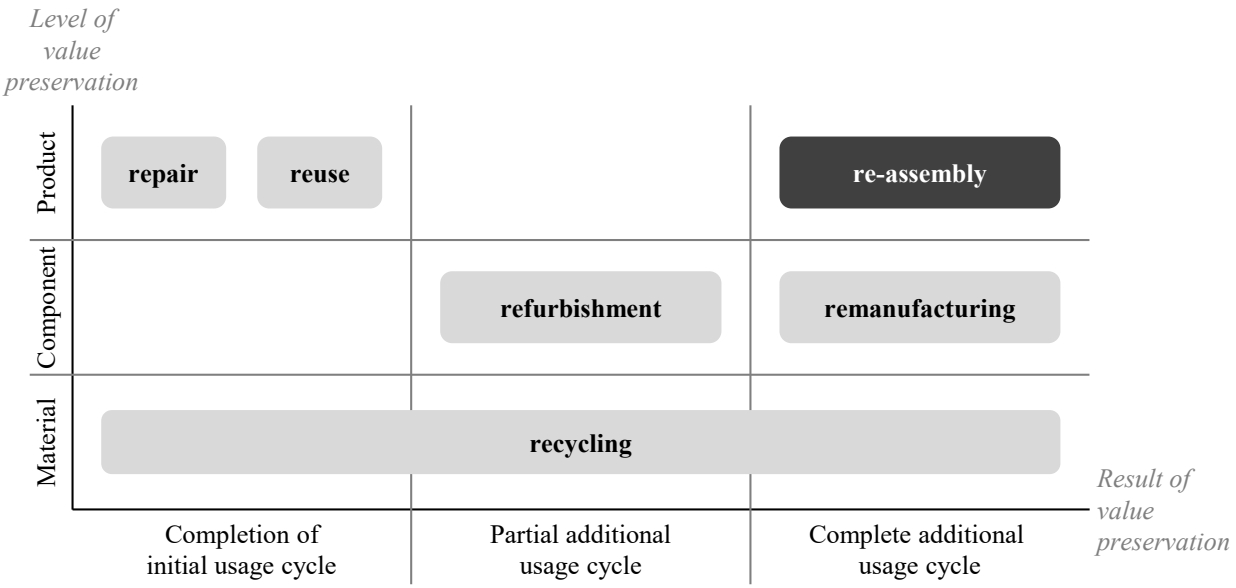


Figure 1: Categorization of re-assembly in terms of level and result of value preservation [8]

In terms of value preservation, Re-Assembly differs in that an entire product is industrially processed rather than just individual components or even just the materials. The result of value preservation is a complete additional usage cycle.

The circular economy itself, and thus also the transition to it, is accompanied by inherent challenges. The inherent uncertainties are the quantity of products, the timing of when these products need to be processed, and the quality of the products to be processed. [9–12] These challenges induce additional challenges, particularly in factory planning and production system design. In order to process products cost-effectively, the production system must be continuously adapted, both qualitatively and quantitatively, to the highly uncertain capacity requirements. [13,14]

2. Theoretical background

2.1 Dimensioning as planning element in factory planning

The just mentioned capacity requirements are part of the dimensioning of factories [15]. The dimensioning is a fundamental planning element in factory planning and the corresponding production system [15,16]. The result of dimensioning is a qualitative and quantitative description of resource requirements for equipment, personnel, and space for a future planning period [17]. The dimensioning process is divided into four sub-steps: production program planning, capacity planning, production segmentation, and determination of organizational forms [15].

- **Production program planning** determines the product range to be sold. The type, quantity, and timing fo the products to be manufactured are specified for defined planning period [15]. In the context of circular economy, it is not only the sales volume that represents the boundary condition, but also the basic availability of cores in a desired quantity, quality, and time [18].
- **Capacity planning** is characterized by the comparison of capacity requirements (demand) and capacity availability (supply). The result of this comparison represents the quantitative and qualitative capacities required to produce the defined production program. Quantitative capacities refer to quantitative performance. Qualitative capacities refer to technical performance. [19]
- **Production segmentation** aims to create fully or partially autonomous units within production system, which bring cost and productivity advantages. This is accompanied by the assumption that coordination within the segments is less complex. Furthermore, production segmentation is based

on the principle of group technology, in which technologies for the product manufacturing process are combined in a single segment. [20–22]

- **Determination of organizational** forms assigns suitable organizational forms to the generated segments. The organizational form characterizes the temporal order of the manufacturing and assembly processes and the spatial order of the production resources. [15,23]

The basic approaches to dimensioning can be divided into static dimensioning and dynamic dimensioning. Statis dimensioning is based on demand for annual periods. It therefore represents an average value. Any dynamics within periods are not considered. This approach has proven to be sufficient in industrial practice in the linear economy. Dynamic dimensioning is based on time-dependent variability in demand volumes. This can be caused by seasonal fluctuations or new product launches. [16] KOBYLKA presents three approaches to capacity planning for fluctuating demand patterns: [24]

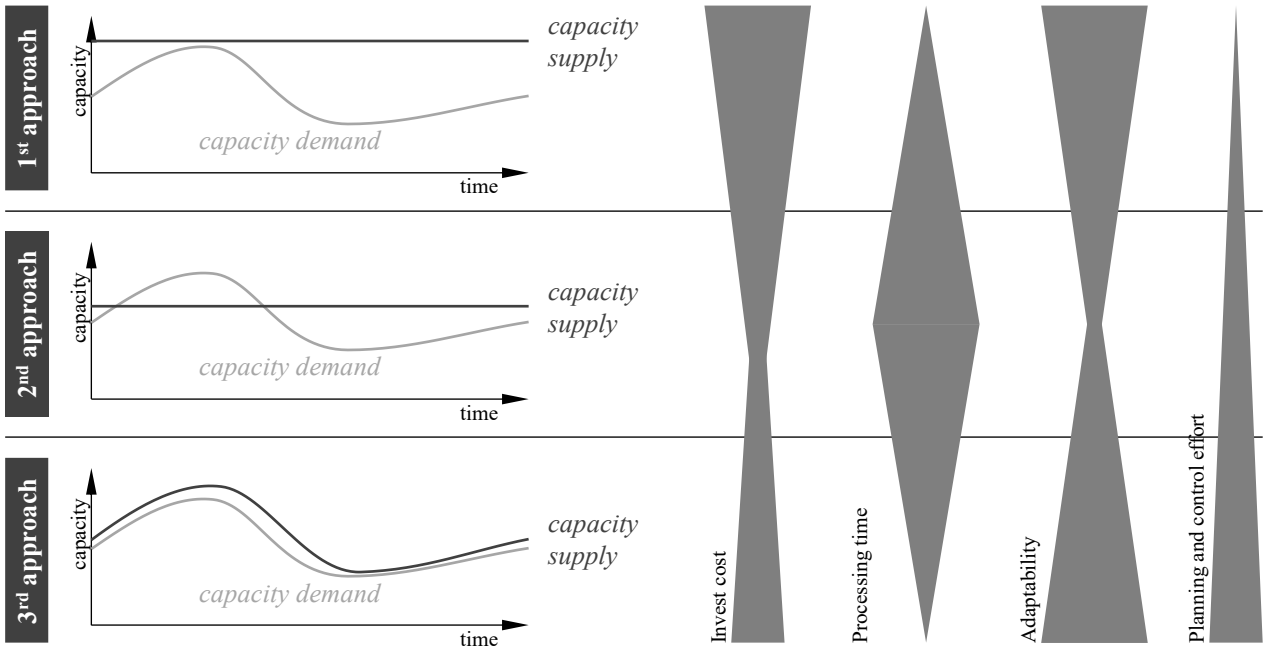


Figure 2: approaches to capacity planning for fluctuating demand patterns [24]

The first approach enables a short processing time, as all capacity is permanently available. This permanent availability results in low planning and control effort. However, this high availability is associated with high investment costs. [24]

The second approach requires smoothing the required capacity. Such capacity shifts lead to increased processing time and associated increased planning and control effort. However, the need for resources is reduced, which means that less investment is required. [24]

The third approach combines the low investment costs of the second approach with the short processing time of the first approach. To implement this, however, capacity requirements must be forecast and capacity in the production system must be continuously activated and deactivated. As a result, the planning and control effort required is greatest. [24]

2.2 Requirements analysis

As circular economy is characterized by a high degree of uncertainty regarding the quantity of products, the timing of when these products need to be processed, and the quality of the products to be processed, the production system must exhibit a high adaptability in order to meet the volatile requirements of re-assembly. [14] In this context, the following four requirements have been recognized:

- Dynamic quantitative and qualitative capacity requirements must be considered so that capacity supply can be implemented in line with demand [25].
- The organizational and spatial structure of the factory must be highly adaptable so that high-frequency and high-amplitude changes in capacity requirements can be met [26].
- The product life cycle data stored in a digital product file must be included in the dimensioning of a factory to reduce the inherent uncertainties of reassembly [27].
- A methodology is needed that proactively designs the dimensioning of a re-assembly factory, considering a digital product file, to provide the capacity required at the time of demand [28].

The development of sustainable value creation factories relies on addressing these requirements through appropriate design and operation of the production system. To adapt the production system accordingly, factory planning research provides relevant approaches, which are discussed in the following.

3. Related work

This chapter presents recent work and contributions in the field of factory planning to the dimensioning of factories in the circular economy.

JURCZYK-BUNKOWSKA [29] proposes an approach to tactical capacity planning using discrete-event simulation combined with throughput cost accounting. The goal is to evaluate capacity expansion or contraction from an economic perspective over a tactical planning horizon. Two key decisions underlie capacity planning: adjusting capacity and selecting the product portfolio, with the author focusing solely on capacity decisions. The discrete-event simulation supports the analysis of the production system, including flexible configuration of modular assembly systems, matching capacity to fluctuating demand, productivity analysis depending on automation level, and decision support for complex processes. When demand exceeds available capacity, a predefined sequence of measures is considered: aligning the production flow to the bottleneck, resolving the bottleneck using lean methods, introducing outsourcing, and implementing incremental process innovations. These measures are simulated in a case study and evaluated by output per time unit. Finally, an economic assessment identifies the most cost-effective option.

GOODAL ET AL. [30] present a framework for simulating material flows in remanufacturing processes. The goal is to predict material flow behavior using data from the production system. The framework consists of three elements: adaptive remanufacturing simulation, a remanufacturing information model, and an information service layer. The remanufacturing simulation model represents the remanufacturing processes within the factory. It is implemented as a discrete-event simulation to determine waiting and processing times along a generic remanufacturing process. The required factory resources are also identified at the start of each run. The validation shows that the production system can respond to changes in resources, capacity requirements, and work content.

KAMPKER ET AL. [11] highlight the potential of remanufacturing battery-electric vehicles. They present a generic remanufacturing process and related scenarios, which are evaluated and compared using the Total Cost of Ownership (TCO) method. The cost assessment consists of the following elements: vehicle acquisition costs, operating costs during use, remanufacturing costs, sales and repurchase costs, and the vehicle's residual value. The authors first classify remanufacturing costs into direct, indirect, fixed, and variable costs. In addition, these costs are assigned to the categories personnel, materials, equipment, and logistics. Within the case study, components are also classified according to their suitability for remanufacturing, distinguishing between high, limited, and no suitability.

ZHANG ET AL. [31] present a two-phase framework for dynamic virtual cell formation (DVCF). The aim is to plan capacity in virtual, dynamically changing manufacturing cells under fluctuating demand across multiple planning cycles. The cells are reconfigured periodically based on demand, where "logical" refers

to high similarity of manufacturing tasks in the sense of group technology. Building on the cell formation problem (CFP), resources are not physically rearranged but virtually grouped into cells.

Each planning cycle consists of several planning periods. The cycle provides the demand-driven capacity data, while the periods are used to reconfigure the virtual cells. The framework defines objectives for each phase: Phase one maximizes process similarity and balances machine workloads. Phase two increases reconfiguration stability by minimizing changes between planning periods. This results in an NP-hard optimization problem, solved using a hybrid Lévy–NSGA-II algorithm to find both locally and globally optimal solutions.

The contributions presented from the literature each include solutions for specific areas in the dimensioning of factories in the circular economy. Nevertheless, none of the identified approaches fully meets all requirements.

4. Methodology for dynamic dimensioning of re-assembly factories

This paper presents a methodology for dynamic dimensioning of re-assembly factories enabling the circular economy. The proposed methodology leverages data from digital product passports to significantly reduce the inherent uncertainty in re-assembly-processes. With the data analyzed, the production system is proactively adapted to fluctuating demand. To ensure that this adaptation is carried out as efficiently as possible, the production system is segmented from a technical and organization perspective. Figure 3 illustrates the methodology's context within the circular value creation system. Figure 4 details the methodology's solution modules.

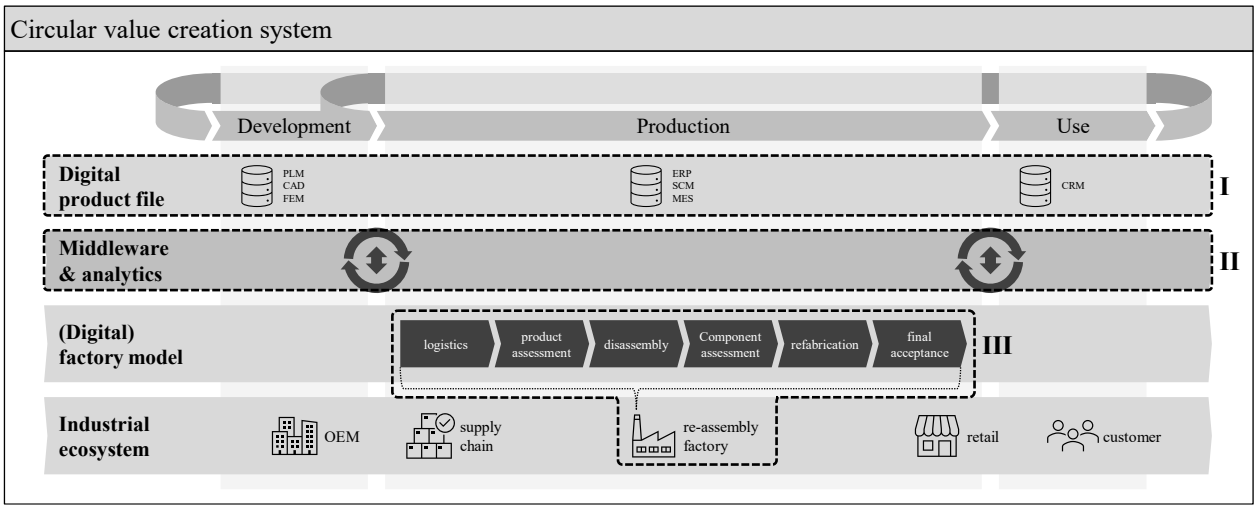


Figure 3: Methodology´s context within the circular value creation system

The methodology is embedded in a circular value creation system, where the decisive temporal dimension is the product life cycle. For the purposes of this work, the product life cycle is simplified and assumed to be quasi-infinite, although a share of primary production and a share of recovery in lower R-strategies remain. The life cycle starts in product development, followed by production, subsequent use, and, after the use phase, a recurring production step in the form of re-assembly.

Vertically, the methodology is structured into four layers. Originating from the product life cycle, the digital product file is positioned as a central element. It consolidates all data generated over the entire life cycle of each product and serves as the single source of truth for all value-adding activities related to the product.

In the context of the Internet of Production, middleware serves as an integration layer that connects heterogeneous data sources, production systems, and IT applications via standardized interfaces and interoperable data models. It enables real-time data acquisition, aggregation, and provision for analytics and

optimization services, thereby decoupling domain-specific applications from the underlying technical infrastructure. [32]

Finally, at the (digital) factory model level, the re-assembly production system is described. This description is based on the set of elements *m*, the processes *p*, and the structure *s*. The elements *m* represent the techno-organizational segments of the production system. The process *p* is represented by the re-assembly process. The structure *s* describes the relations between the elements *m*. [33]

The industrial ecosystem level describes the actors in the real value creation system. These actors include the original equipment manufacturer (OEM), the supply chain, the re-assembly-factory, retail, and the customers. The actors interact both through data and through material flows. It should be noted that different configurations of actors and interaction patterns are possible. However, the relevant scope of analysis in this work is limited to the production site. [34]

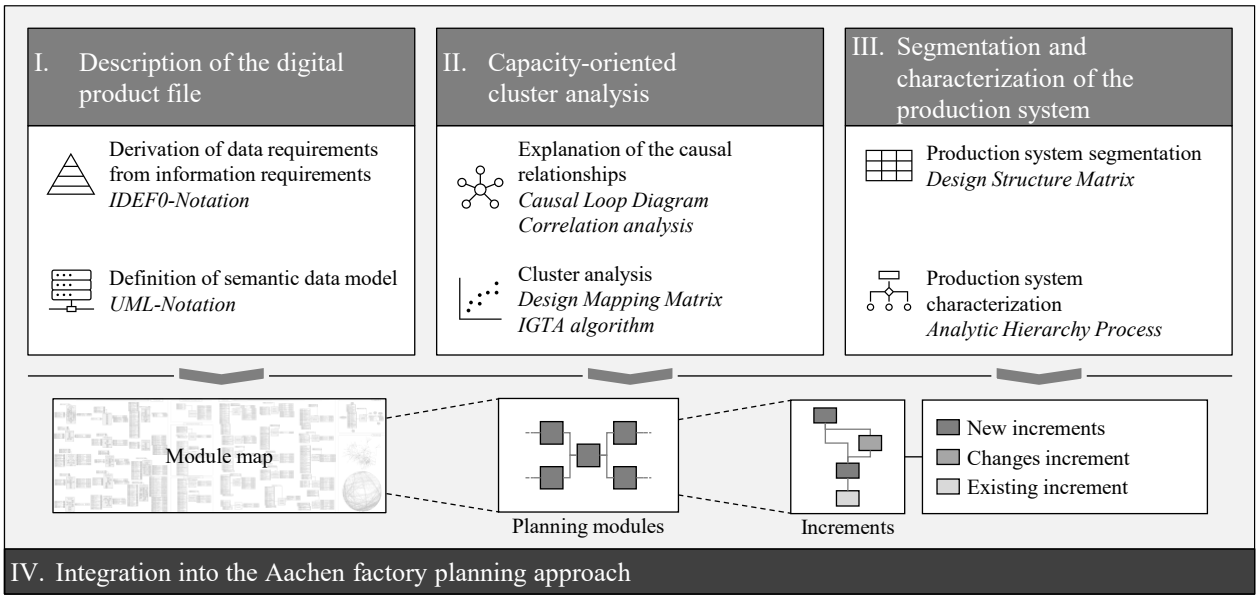


Figure 4: Methodology's solution modules

In total, the methodology comprises four solution modules. Solution modules I to III are derived from the requirements of the circular value creation system. Solution module IV serves to formalize and systematize the methodology in the context of factory planning. The four modules are explained in the following.

4.1 Description of the digital product file

The purpose of this module is to describe the required data needs within a digital product passport. To this end, the necessary information requirements for dynamic dimensioning are first derived. This derivation is modeled using IDEF0 notation. A distinction is made between product-specific data, production-specific data, and usage-specific data. Product-specific data are defined during product development and remain constant over the entire life cycle. Production-specific data are defined in the context of primary production and re-assembly and have a limited validity. Usage-specific data originate from the usage cycles and likewise exhibit a limited validity. The result is a semantic data model specified in UML.

4.2 Capacity-oriented cluster analysis

The capacity-oriented cluster analysis links the data of the digital product passport with the data requirements of capacity planning. For this purpose, qualitative descriptions and quantitative explanations of the causal relationships between the data points are developed using a causal loop diagram. The diagram describes and explains the required demands for labor, space, and equipment in several interrelated contexts with decreasing temporal weighting. To perform the subsequent cluster analysis, a design mapping matrix is

constructed. Vertically, the re-assembly process is described, while horizontally the products considered within a given time slice are represented. Additional matrices with an identical structure are defined for subsequent time periods. An adapted IGTA algorithm then analyzes clusters in three dimensions, optimized for a minimal number of clusters and a minimal number of required changes.

4.3 Segmentation and characterization of the production system

The results of the cluster analysis are evaluated, and a design structure matrix is used to describe the segmentation of the production system for a given planning period. Subsequently, the segments are characterized according to organizational principles, taking into account aspects such as the volume and variety of material flows, the required flexibility of work content, and intralogistics requirements. Using an analytic hierarchy process, the segments are then classified as line, matrix, job shop, group, or cellular manufacturing.

4.4 Integration into the Aachen factory planning approach

To formalize these modules, the promising Aachen factory planning approach is applied. This approach represents a modular planning concept that provides the required flexibility and adaptability. It systematically encapsulates the planning content for the production system and the industrial building in planning modules, each of which comprises more finely structured planning increments. For the formalization of the methodology for the dynamic dimensioning of re-assembly factories, increments from the planning modules capacity planning, production structure planning, process planning, product analysis, and production program analysis are considered and, where necessary, adapted. In doing so, existing increments are reused, some increments are modified, and additional increments are newly introduced.

5. Test environment

The validation of the proposed methodology will take place within the DiCES research project. The project partners are *Miele & Cie KG, the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, All for One Group SE, Forcam GmbH, IconPro GmbH, Klima.Metrix GmbH, and FIR at RWTH Aachen University.* The project´s goal is to develop an innovative, data-based value chain system that seamlessly integrated circular economy principles into companies´ operations. [35]

The use case considers original equipment manufacturer (OEM) that produces white goods with a degree of in-house manufacturing. The manufacturer adopts adjusted variants of the re-assembly-process named refurbishment light, refurbishment and remanufacturing. [14] The initial product assessment is carried out at decentralized locations to optimize the re-assembly-process. The existing linear production system is expanded by additional areas to handle re-assembly processes within the same factory. [34]

Washing machines are well suited to re-assembly due to their modular design and durable components such as drums, motors, and housings. At the same time, the white goods sector faces specific challenges, including high wear of components such as rubber seals, electronic controls, and pumps, as well as long service life and typically high sales price. In addition, reverse logistics for end-of-life products is complex and cost-intensive, particularly in geographically dispersed markets, adding uncertainty and operational complexity to circular production planning.

The validation focuses on the dimensioning of the production system. The emphasis is not on layout design or on the well-known challenges related to control and intralogistics. However, the high diversity and variability of product variants, conditions, and refurbishment requirements make this case particularly suitable for the intended use case. The validation is structured into two strands. Strand 1 focuses on a dimensioning approach that corresponds to the practices predominantly established in industrial practice, while Strand 2 focuses on the dynamic dimensioning approach proposed in this work.

6. Summary and outlook

The implementation of the circular economy is beneficial from both an economic and an ecological perspective. Within this paradigm shift, re-assembly offers an innovative approach to continue providing products with new-product quality and regular upgrades to the market, without simultaneously increasing the consumption of primary resources. However, this also requires the availability of suitable production facilities for re-assembly-based refurbishment.

This work presents a methodology that supports the dynamic dimensioning of re-assembly factories. The dynamic dimensioning follows the volatile demand that is induced by the inherent uncertainties in the circular economy. The methodology leverages data from the digital product passport to forecast the capacity requirements for future planning periods. Building on this, a proactive adaptation of the production system is carried out, which is facilitated by a segmented techno-organizational structure.

Further research is needed regarding validation for additional products and industries, the sensitivity to the data available in the digital product passport, and the transfer to other value creation concepts e.g., contract-based remanufacturing.

Acknowledgements

Funded by the Federal Ministry of Economic Affairs and Climate Action (BMWK) in the GreenTech funding framework “Development of digital technologies” as the joint research project DiCES (Digital Transformation of Circular Economy for Industrial Sustainability) with the funding code 01MN23022E.

References

[1] Ellen MacArthur Foundation, 2013. Towards The Circular Economy: Economic and business rationale for accelerated transition.

[2] Dobbs, R., Oppenheim, J., Thompson, F., Brinkman, M., Zornes, M., 2011. Resource Revolution: Meeting the world´s energy, materials, food, and water needs.

[3] European Commission, 2020. Aktionsplan für die Kreislaufwirtschaft: Für ein saubereres und wettbewerbsfähigeres Europa. https://environment.ec.europa.eu/strategy/circular-economy-action-plan_en?prefLang=de. Accessed 14 May 2024.

[4] DIN Deutsches Insitut für Normung e.V., DKE Deutsche Kommission Elektrotechnik, VDI Verein Deutscher Ingenieure e.V., 2023. Deutsche Normungsroadmap Circular Economy, Berlin. https://www.din.de/resource/blob/892606/06b0b608640aaddd63e5dae105ca77d8/normungsroadmap-circular-economy-data.pdf.

[5] Suzanne, E., Absi, N., Borodin, V., 2020. Towards circular economy in production planning: Challenges and opportunities. European Journal of Operational Research 287 (1), 168–190.

[6] Potting, J., Hekkert, M., Worrell, E., Hanemaaijer, A., 2017. Circular Economy: Measuring innovation in the product chain.

[7] Riesener, M., Schmitz, S., Fulterer, J., Janssen, B., Obladen, A., Kocyigit, E., Pomykol, P., Schuh, G., 2024. Circular Economy in der Produktion: Re-Assembly und Recycling, in: Zäh, M. (Ed.), Handbuch Nachhaltige Produktion. Rahmenbedingungen, Werkzeuge, Anwendungsfelder. Hanser, München, pp. 259–286.

[8] DIN Deutsches Institut für Normung e.V., 2023. Remanufacturing (Reman): Qualitätsklassifizeirung für zirkuläre Prozesse. Beuth Verlag GmbH, Berlin.

[9] Guide, V.R., 2000. Production planning and control for remanufacturing: industry practice and research needs. Journal of Operations Management 18 (4), 467–483.

[10] Hollah, A., 2020. Planung und Steuerung des Remanufacturing in der Elektromobilproduktion: Remanufacturing planning and control in electromobility production, 1. Auflage ed. Apprimus Verlag, Aachen, 206 pp.

[11] Kampker, A., Hollah, A., Triebs, J., Aidi, I., 2019. Remanufacturing of electric vehicles: Evaluation of economic potential, in: 2019 IEEE 10th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT). 2019 IEEE 10th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), Cape Town, South Africa. 15.02.2019 - 17.02.2019. IEEE, pp. 122–127.

[12] Kurilova-Palisaitiene, J., Sundin, E., Poksinska, B., 2018. Remanufacturing challenges and possible lean improvements. Journal of Cleaner Production 172, 3225–3236.

[13] Bonvosin, J., Stark, R., Seliger, G., 2017. Field of Research in Sustainable Manufacturing, in: Stark, R., Seliger, G., Bonvoisin, J. (Eds.), Sustainable manufacturing. Challenges, solutions and implementation perspectives. Springer Open, Cham, pp. 3–20.

[14] Schuh, G., Schukat, E., Schäfer, N., 2025. How to Design Production Systems for Circular Economy: Applying the Modular Aachen Factory Planning Approach to a White Goods Manufacturer, in: International Conferen on Remanufacturing. International Conference on Remanufacturing (ICoR 2025), Amsterdam. 08.-10. April.

[15] Burggräf, P., Schuh, G. (Eds.), 2021. Fabrikplanung: Handbuch Produktion und Management 4, 2nd ed. ed. Springer Berlin / Heidelberg, Berlin, Heidelberg, 566 pp.

[16] Schenk, M., Wirth, S., Müller, E., 2009. Factory Planning Manual. Springer, Dordrecht, 421 pp.

[17] Grundig, C.-G., 2021. Fabrikplanung: Planungssystematik – Methoden – Anwendungen, 7., aktualisierte Auflage ed. Hanser, München, 333 pp.

[18] Tolio, T., Bernard, A., Colledani, M., Kara, S., Seliger, G., Duflou, J., Battaia, O., Takata, S., 2017. Design, management and control of demanufacturing and remanufacturing systems. CIRP Annals 66 (2), 585–609.

[19] Eversheim, W., 1989. Organisation in der Produktionstechnik, 2., neubearb. u. erw. Aufl. ed. VDI-Verl., Düsseldorf, 335 pp.

[20] Bergholz, M., 2005. Objektorientierte Fabrikplanung.

[21] Eversheim, W., Neuhausen, J., 2001. Modular Plant Architecture. wt Werkstattstechnik online 91 (10), 654–657.

[22] Wildemann, H., 1994. Die modulare Fabrik: Kundennahe Produktion durch Fertigungssegmentierung, 4., neubearb. Aufl. ed. TCW-Transfer-Centrum, München, 421 pp.

[23] Reuter, C., Prote, J.P., Gützlaff, A., Luckert, M., 2016. Generation of Cellular Manufacturing Variants for the Potential Determination of Production Structures. AMR 1140, 489–496.

[24] Kobylka, A., 2000. Simulationsbasierte Dimensionierung von Produktionssystemen mit definiertem Potential an Leistungsflexibilität. Dissertation, Chemnitz.

[25] Huster, S., Rosenberg, S., Glöser-Chahoud, S., Schultmann, F., 2023. Remanufacturing capacity planning in new markets—effects of different forecasting assumptions on remanufacturing capacity planning for electric vehicle batteries. Jnl Remanufactur 13 (3), 283–304.

[26] Eisbein, H., Riesener, M., Schmitz, S., Fulterer, J., 2025. Production Structures for Circular Production: A Systematic Literature Review, in: Drossel, W.-G., Ihlenfeldt, S., Dix, M. (Eds.), Production at the Leading Edge of Technology. Springer Nature Switzerland, Cham, pp. 598–607.

[27] Schuh, G., Mauß, W., Potente, T., Schmitz, S., Adlon, T., Maetschke, J., Neumann, H., Salzwedel, J., Kozielski, S., Luckert, M., Reuter, C., Schmidhuber, M., Witthöft, J., 2023. Green Re-Assembly Upgrade Factory, 314-338.

[28] Schuh, G., Schukat, E., Lehde genannt Kettler, N., Danwitz, F. von, 2025. Fabrikplanung für die Re-Assembly/Factory design for re-assembly. wt 115 (01-02), 17–25.

[29] Jurczyk-Bunkowska, M., 2021. Tactical manufacturing capacity planning based on discrete event simulation and throughput accounting: A case study of medium sized production enterprise. Adv. Prod. Eng. Manag. 16 (3), 335–347.

[30] Goodall, P., Sharpe, R., West, A., 2019. A data-driven simulation to support remanufacturing operations. Computers in Industry 105, 48–60.

[31] Zhang, Z., Guan, Z., Fang, W., Yue, L., 2023. DYNAMIC VIRTUAL CELLULAR RECONFIGURATION FOR CAPACITY PLANNING OF MARKET-ORIENTED PRODUCTION SYSTEMS. J. Ind. Manage. Optim. 19 (3), 1611–1635.

[32] Schuh, G., Prote, J.-P., Gützlaff, A., Thomas, K., Sauermann, F., Rodemann, N., 2019. Internet of Production: Rethinking production management, in: Wulfsberg, J.P., Hintze, W., Behrens, B.-A. (Eds.), Production at the leading edge of technology. Proceedings of the 9th Congress of the German Academic Association for Production Technology (WGP), September 30th - October 2nd, Hamburg 2019. Springer, Berlin, Heidelberg, pp. 533–542.

[33] Schmigalla, H., 1995. Fabrikplanung: Begriffe und Zusammenhänge. : im Auftrag des VDA-AK Arbeitsgruppe "Rechnerungterstütze Fabrikplanung". Carl Hanser, München, 407 str.

[34] Perau, M., Gaillard, A., Spindler, D., Schuldt, F., Özkan, C., Schröer, T., Boos, W., 2024. Value Systems for the Parallel Implementation of Value Retention Circular Strategies in the White Goods Industry, in: PRO-VE 2024: Navigating Unpredictability: Collaborative Networks in Non-linear Worlds. PRO-VE 2024: Navigating Unpredictability: Collaborative Networks in Non-linear Worlds, Albi, France. 24.10.2024-30.10.2024, pp. 351–368.

[35] Gaillard, A., Perau, M., 2023. DiCES: Digital Transformation of Circular Economy for Industrial Sustainability: Von linearem zun zirkulärem Wertschöpfungssystem: Implementierung von ReX-Strategien in die Auftragsabwicklung. UdZ - The Data-driven Enterprise 3 (1), 93–97.

Biography

Günther Schuh (*1958) holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering (WZL) at RWTH Aachen University, is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT and director of the Research Institute for Rationalization (FIR) at the RWTH Aachen.

Esben Schukat (*1993) is head of Factory Planning department at the Laboratory for Machine Tools and Production Engineering (WZL) at RWTH Aachen University since 2023. He is also managing director of the Center Factory Planning and the Center for Smart Industrial Agriculture as well as Senior Manager of the Center Smart Assembly.

Nils Lehde genannt Kettler (*1992) studied mechanical and production engineering at the RWTH Aachen University. He is the team leader of Digital Factory Planning and research associate at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. He is also senior consultant at the Center Factory Planning.

8th Conference on Production Systems and Logistics

Synergies Among Urban Mining, Reverse Logistics, And Closed-Loop Supply Chains Towards Circular Economy: A Decision Support Framework

Kuruppuge Themiya, Kulatunga Asela, Luis Martino, Sucala Voicu Ion

Exeter Digital Enterprise (ExDES) Lab, Department of Engineering, Faculty of Environment, Science and Economics, University of Exeter, Exeter, UK

Abstract

Although several concepts have emerged over time that support the circular economy (CE) either directly or indirectly, it is rare to find any studies that try to use them as an integrated approach or even to complement each other towards the final goal. Even though there are some overlaps of these concepts, which address the end-of-life (EoL) phase of products, there is a notable lack of studies that attempt to integrate each of these concepts by considering the stakeholder concerns and the ground-level realities promoting the circular economy. However, synergies of product EoL concepts could contribute toward the CE to a greater extent. Therefore, the present study examines how reverse logistics (RL), urban mining (UM), and closing the loop of the supply chain can be integrated to support the circular economy. The study proposes an integrated decision support framework, identifies different options and avenues available towards the recovery process, and proposes different decisions at various levels of the EoL stage of the product towards each stakeholder. Framework effectiveness is validated through the Material Circularity Indicator analysis of the washing machine recovery process, achieving circularity higher than isolated implementations.

Keywords

Decision Support Framework; Urban mining; Closed loop supply chain; Product End-of-life management; Material recovery

1. Introduction

The CE seeks to capture the economic and material value lost within traditional linear systems. In 2023, more than 90% of materials were lost during compliant recycling processes or disposed of through ineffective, non-compliant processes and dumpsites [1]. Considering the unregulated waste dumping, pollution, and long-term waste-management issues, transitioning to CE systems has become imperative, compared to the traditional take-make-dispose model. Several concepts have emerged to facilitate the goals of the circular economy, notably UM and closed-loop supply chains (CLSC). All these strategies aim at increasing the lifespan of materials and products, maximising the operational use of the materials, and finally recycling the materials at the end of the life cycle of a product. UM involves the recovery of useful materials from human-made stocks and waste streams, particularly in urban areas where discarded products accumulate to economically viable concentrations. CLSC extend traditional supply chain management to include product returns and material recovery, which creates a cycle of products that re-link EoL products to the manufacturing process, where feasible [2].

Although these approaches share common objectives in advancing the circular economy, they have evolved independently. Integration between them is still rare [3,4]. UM supports the extraction of valuable resources

DOI: https://doi.org/10.15488/21019
ISSN: 2701-6277

by organising flows of materials, information, and service networks, thus making it possible to extract the valuable resources of various anthropogenic sources [5]. Unlike traditional mining, which extracts virgin resources, UM recovers materials from discarded electronics, vehicles, and other consumer goods [6]. However, it mainly focuses on material recovery with limited attention on the collection infrastructure needed to access materials. RL can address this gap by providing the operational framework for collection, directing resources extracted through UM back into the supply chain, and creating closed-loop material flows. It addresses the planning, executing, and managing the effective flow of products, parts, and materials out of consumer locations to points of origin or proper disposal [7]. While this encompasses EoL products, its combination with UM is limited in the literature and practice [4].

Addressing these challenges requires more than simply combining existing concepts. Without integration, valuable materials may be found but never recovered, collection systems may operate without understanding their full value potential, and system-level goals may remain disconnected from operational realities. It demands structured approaches to guide stakeholders through decisions at multiple levels. Developing a decision-support tool can cover the full range of EoL product management. As a result, this study develops a decision support framework synthesising the idea of UM, RL, and CLSC to support the practical implementation of CE practices. The proposed framework provides structured guidance for EoL product management decisions across multiple recovery pathways while addressing the distinct concerns of different stakeholders in the system. It defines the decision criteria to be used during the selection of routes at distinct phases of the recovery process. The remainder of this paper will be structured as follows. Section 2 reviews existing frameworks and identifies integration gaps. Section 3 describes the methodology. Integrated framework architecture, with decision hierarchies and the stakeholder coordination mechanisms presented in Section 4. Finally, Section 5 provides the discussion of the findings and concluding remarks.

2. Literature Review

Post-usage management is a critical lever for CE implementation, as EoL strategies determine whether material value is retained, degraded, or lost completely from the economic cycle [8]. Within post-usage pathways, UM, RL, and CLSC play a pivotal role. UM has become one of the most critical approaches towards the recovery of high-value materials. Ottoni et al. [9] illustrate how UM and RL can be operationalised, yet the study concludes at the recycling facility gate, failing to close the circular loop. Xavier et al. [5] examine UM from a broader perspective, analysing diverse definitions of value recovery and systematically reviewing the stages of e-waste UM, including the exploitation of anthropogenic resources as secondary sources of critical raw materials such as Au, Ag, Cu, Li, and Co. Nevertheless, their review does not provide operational frameworks to bridge material recovery with collection infrastructure and manufacturing reintegration pathways. Similarly, Simaei et al. [10] suggest a multi-criteria decision support system to evaluate the product-level recovery options. However, the system operates in isolation without considering the logistics required to collect products or the CLSC mechanisms needed to reintegrate recovered materials into manufacturing. Freitas et al. [11] examine reverse supply chains, UM processes, and CE principles through system dynamics, demonstrating that tailored regulations can improve resource recovery efficiency. Yet it remains at a conceptual strategic level, providing policy recommendations without translating these insights into operational integration frameworks. The summary given in Table 1 was developed to identify which frameworks cover UM, RL, and CLSC independently or in combination with recovery pathways based on several criteria: (1) the breadth and depth of the CE principles; (2) the extent of integration between material recovery as well as reverse-logistics flows; (3) the decision-making processes at EoL management; (4) the stakeholder considerations; and (5) the recovery pathways (R-strategies, which encompass reuse, rethink, reduce, repair, refurbish, remanufacture, repurpose, recycle, and recovery) described. Existing frameworks fail to integrate UM, RL, and CLSC operationally. This disintegration prevents the result of synergies where material expertise is used as an informational source to determine collection priority, logistics capacities to support scalable processing, and manufacturing systems generate

demand for recovered materials. The remainder of this paper addresses these gaps through an integrated operational framework coordinating these elements.

Table 1: Summary of previous frameworks

Study	Urban Mining Focus	Reverse Logistics Integration	CLSC Consideration	Recovery Strategies	Limitations
System dynamics applied to the e-waste value chain: A Brazilian case study [11]	✓	✓	✓	R3, R6, R8	No operational framework and multi-stakeholder coordination
A circular approach to the e-waste valorization through urban mining in Rio de Janeiro, Brazil [9]	✓	✓		R8	No closed-loop reintegration
Improving urban mining practices for optimal recovery of resources from e-waste [12]	✓			R3, R8, R9	Focused only on material recovery processes
AI-based decision support system for enhancing end-of-life value recovery from e-wastes [10]	✓			R8	No collection infrastructure and material reintegration
Sustainable electronic waste management framework for academic institutions in India [13]	✓	✓		R4, R8	Focused on E-waste management in an academic institution; No material reintegration and multi-stakeholder coordination
A systematic review of the factors – Barriers, drivers, and technologies – Affecting e-waste urban mining: On the circular economy future of developing countries [14]	✓		✓	R8	Focused on urban mining stages and essential practices; No operational framework
A comprehensive review of urban mining and the value recovery from e-waste materials [5]	✓		✓	R4, R5, R6, R8	No operational framework

Note: R0 Refuse; R1 Rethink; R2 Reduce; R3 Reuse; R4 Repair; R5 Refurbish; R6 Remanufacture; R7 Repurpose; R8 Recycle; R9 Recover

3. Methodology

This study employed a four-phase research approach comprising a literature review, framework development, data collection and analysis, and conclusion. As presented in Figure 1, the initial literature review was conducted to identify interchangeable terms for the primary keywords used in the database search. Boolean operators (“AND” and “OR”) were used to construct the search query based on the terms “urban mining”, “supply chain”, “circular economy”, and “reverse logistics”. The Scopus keyword search was conducted using the TITLE-ABSTRACT-KEYWORDS fields. The search was limited to English-language documents and included only articles, reviews, conference papers, and book chapters. The time span was set to 2000–2025, and 55 documents were initially identified for manual screening. After manual screening, 16 documents were selected. The framework was developed and refined based on insights from the literature review. Key decision nodes across the EoL lifecycle were identified, and based on these decisions, recovery pathways were outlined at the product, component, and material levels. A stakeholder analysis was conducted to identify key players such as manufacturers, collectors, recyclers, and regulators, and their position in the strategic, tactical, and operational decision-making to enhance circularity. Thinkstep-anz’s Excel-based [15] Material Circularity Indicator (MCI) calculator [15] was used to measure circularity, which was developed using the methodology recommended by the Ellen MacArthur Foundation, which is aligned with ISO 59020:2024 standards [16]. Finally, secondary data from reports, policy

documents, and other organisational documents were used to evaluate the performance of circularity in different recovery pathways.

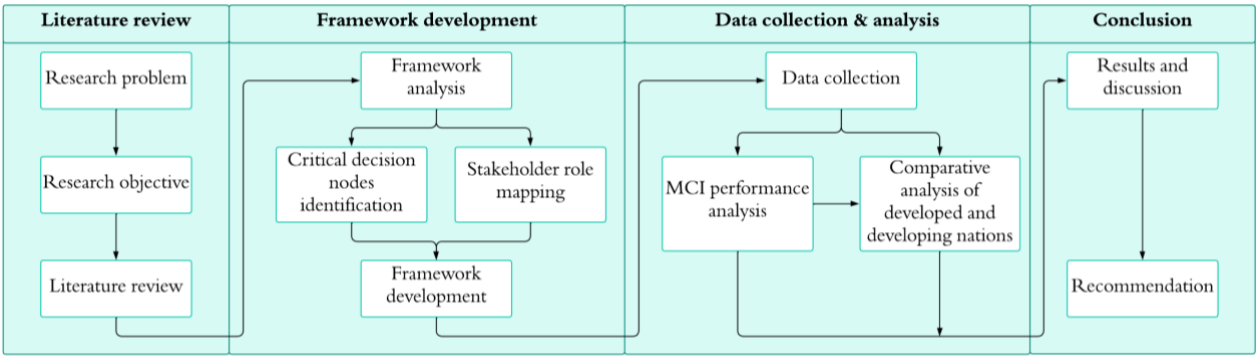


Figure 1: Schematic representation of research methodology

4. Framework Development

4.1 Proposed framework

The key objective of the study is to identify synergies between UM, RL, and the CLSC. To achieve synergies between these three concepts, this study develops an integrated framework through mapping the material flows, stakeholder interactions, and decision hierarchies across the product lifecycle. To establish the integration points, the whole supply chain of the raw material extraction to the EoL was outlined to highlight the areas of integration. Then the reverse-logistics pathways were superimposed onto the forward supply chain to emphasise the opportunity of resource circulation. Finally, UM decision nodes within RL were embedded to determine the recovery routes based on product condition, component value, and material composition. With this approach, the framework captures not merely sequential processes but the interdependencies and feedback loops that generate synergistic effects exceeding individual component capabilities.

As highlighted in Figure 2, UM involves testing, inspection, and recovery processes that identify the right channels through which EoL products should be taken. Critical decision nodes which identified in the literature review [17] are embedded throughout each level to manage waste efficiently without losing value. The framework clearly shows both formal and informal collection paths, recognising the key role that informal players have in various situations in developing economies. This recognition addresses a practical reality often overlooked in frameworks designed primarily for developed economy contexts, where formal collection systems dominate. The framework outlines three main recovery paths that include product recovery, component recovery, and material recovery. All the pathways have different degrees of disassembly and processing intensity, and each of them has its implications on value retention and resource consumption.

Recovery paths are based on the waste hierarchy, with the R-strategies in the first place that conserve the highest value. The most preferred option is product recovery, and the least preferred is landfilling. UM focuses on value recovery from EoL products, yet most applications concentrate on material recycling. This emphasis appears in both academic literature and industrial practice, where recycling infrastructure and expertise have developed more extensively than systems for product recovery or component recovery. The proposed framework expands this concept boundary to capture more product value by moving up the value hierarchy from material recycling to component and product recovery. This gives an aid in determining and extracting value at the product and component levels before moving out to material recycling. The framework criticizes the traditional focus on recycling by showing how recovery strategies of higher order can bring better environmental and economic impact.

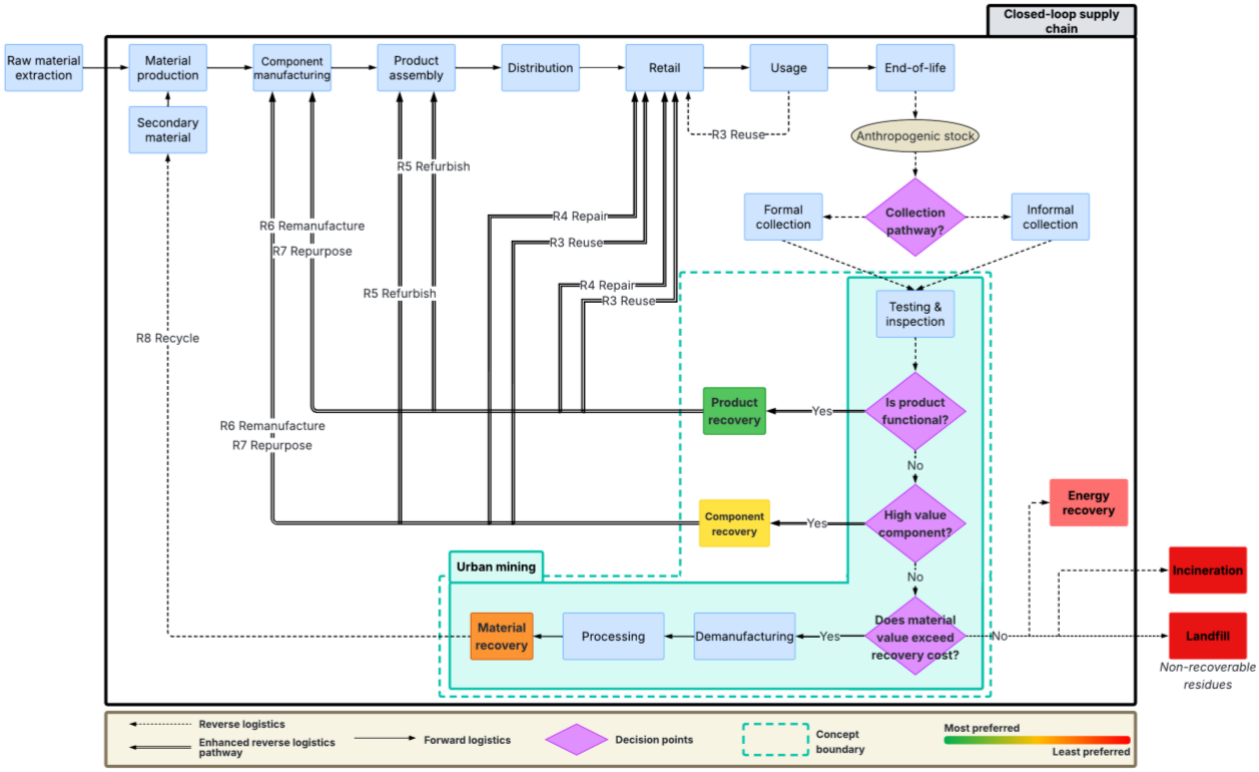


Figure 2: Proposed framework

4.2 Decision levels and stakeholder responsibilities

The management of EoL requires decision-making processes at the strategic, tactical, and operational levels. Each addresses distinct time horizons, resource scales, and stakeholder responsibilities that collectively determine CE performance. Table 2 presents the key decision nodes in the process and shows the role of the stakeholders in placing their positions to manage the waste responsibly. These levels are an interrelated hierarchy, where decisions at one level have effects on opportunities and limitations at the other levels. Strategic decisions are made by manufacturers, policymakers, and waste-management organisations on a long-term basis. Tactical decisions deal with medium-term planning, which involves a connection between strategic infrastructure and operational implementation and translating system-level goals into action plans. Operational decisions govern the day-to-day execution of recovery processes. While more routine than strategic or tactical choices, they directly decide realised system performance. The operational decisions made have a cumulative impact of either making or breaking the system in line with its strategic objectives, despite the good high-level planning. The framework acknowledges that effective EoL management must have a sense of coherence at all three levels. Strategic policies that do not consider operational realities will not be effective, and operational efficiency cannot be achieved at the expense of strategic policy and tactical failure. The multi-level view helps stakeholders understand how their actions relate to the overall system outcomes and where they need to coordinate with other actors.

Table 2: Decision levels and stakeholder mapping

Decision Node	Decision Level	Decision	PR	OEMs	MC	3rd Party Service Providers		
						RC	CL	W&T
Collection pathway	S	Define integration of formal and informal collection channels. Establish take-back and EPR structures.	✓	✓				
	T	Design collector incentive schemes. Select licensed partners.		✓		✓		
	O	Identification, handling, sorting.			✓		✓	
Functionality assessment	S	Establish standardised inspection guidelines. Safety compliance.	✓	✓				
	T	Allocate testing tools. Train workers, infrastructure, and workflows.				✓		
	O	Visual tests, electrical tests, safety checks.						✓
Component value assessment	S	Prioritise high-value components. Secondary market requirements.		✓				
	T	Determine recovery order.				✓		
	O	Extract reusable modules. Check quality.						✓
Material value assessment	S	Forecast critical material supply risks. Recycling plants' capability.	✓			✓		
	T	Select separation processes (pyro, hydro, mechanical). Define preprocessing layouts.				✓		
	O	Conduct shredding, separation, sampling, and purity checks.						✓

Note: S: Strategic; T: Tactical; O: Operational

PR: Policy Regulators; OEMs: Original Equipment Manufacturers; MC: Municipal Councils; RC: Recyclers; CL: Collectors; W&T: Workers & Technicians (personnel who execute daily material handling and processing tasks)

5. Results and Discussion

5.1 Material circularity performance

Enhancement of circularity was quantified using the MCI calculator [15] within the developed framework, specifically to highlight the synergetic contributions. MCI quantifies circularity performance by measuring the balance between virgin material inputs, material recovery at end-of-life, and product utility through lifespan extension. This was used because of the credibility and based on material flow analysis. A standard household electrical appliance: a 7 kg top/front-loading washing machine, was chosen as a representative case study to illustrate how such a product moves through the different pathways outlined in the proposed framework, thereby contributing to enhanced circularity. The analysis evaluates how the product moves through different recovery pathways across three waste-collection contexts: EU with formal systems, the United Kingdom (UK) with both formal and informal practices, and Sri Lanka (SL) with informal networks and emerging formal structures. The analysis was conducted across three scenarios: product reuse, component recovery, and material recycling. The system boundary considered material extraction through EoL recovery. Seven materials were evaluated using the simplified MCI tool: bulk plastics (PP), technical plastics (nylon), ferrous metals (steel), non-ferrous metals (aluminium), electronics (PCBs), glass, and paper. Baseline assumptions included 100% virgin material inputs and a 12.5-year lifespan [18]. Figure 3 indicates the value distribution across recovery pathways in the EU context [18]. For reuse and remanufacturing pathways, a 100% recovery rate was assumed as an ideal benchmark, while EU recycling rates [18] for the listed materials were applied in the recycling scenario. Product reuse yields a 50 % circularity in all materials, hence keeping the entire product without reworking. The 50% ceiling reflects virgin material assumptions; incorporating reuse or recycled content in the production process would increase circularity beyond the given mark, indicating once again that circularity depends on the sourcing during the production process as well

as EoL recovery. Recovery through remanufacturing of components achieves a circularity rate of 48%, only 2 % lower than that achieved through reuse. The small difference in performance is due to remanufacturing keeping the product's architecture with the selective replacement of the worn parts. Circularity in material recycling ranges from 3% (glass) to 40% (steel/aluminium). Ferrous metals score relatively high on the circularity scale because there is minor degradation of quality, as compared to technical plastics and electronic components that illustrate low recyclability, as they are difficult to separate from the polymer and the heterogeneous material matrix. For electronics, the advantage over recycling becomes dramatic: 48% versus 3%, a 45-point differential reflecting the complexity and material diversity in PCBs that recycling cannot efficiently recover. The circularity performance shows that the best value retention involves a combined implementation of RL, UM, and CLSC. Each addresses distinct aspects of circular value capture, and their synergy decides the overall circularity of the system.

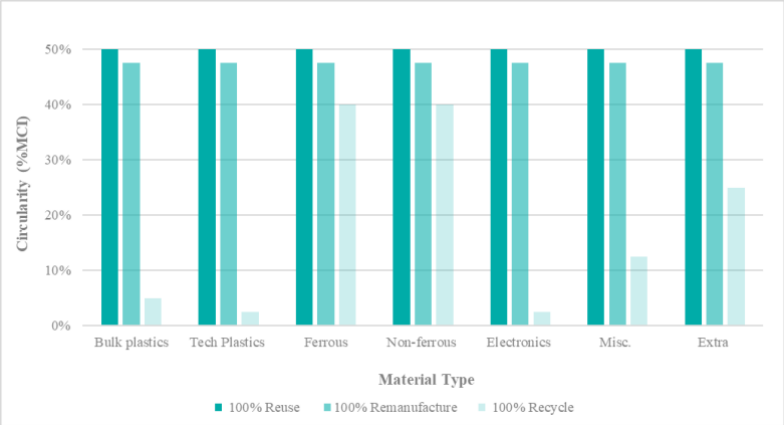


Figure 3: Material circularity performance for the 7kg washing machine base case

5.2 Comparative Analysis: UK vs Sri Lanka Resource Recovery Systems

MCI analysis reveals different circularity pathways between developed and developing economies. Table 3 presents the performance under current operational conditions.

Table 3: Material circularity indicator comparison

Parameter	United Kingdom	Sri Lanka	EU (Spain)
Recycling rate	57% [19]	9% [20]	77% [21]
Reuse rate	2% [22]	NA (Not available)	12% [21]
Average product lifespan	12 Years [23]	12 Years (Assumed)	12 Years (Assumed)
MCI score	32%	7%	52%
MCI score (at 15 years)	46%	26%	61%

These results highlight the contrasting approaches to resource recovery and CE implementation. Figure 4 illustrates the correlation between product lifespan and circularity performance, demonstrating that advanced collection and recycling infrastructure yields compounding returns when combined with extended product life. However, Sri Lanka's formal recycling rate of 9% underestimates actual material recovery within its informal economy. Washing machines contain valuable ferrous, non-ferrous metals, and plastics, creating a strong economic reason for collection, and they rarely reach landfills without material extraction. Figure 5 presents this scenario across various lifespans (11.6, 15, 17.5, and 20 years) with adjusted recycling rates of 10–50%.

UK leads in UM and secondary material recovery compared to SL, based on the recycling rates. Upstream inefficiencies limit success downstream because only a small part of the recoverable materials reach these facilities. Underdeveloped reuse culture compounds these challenges. In the UK, reuse rates remain 2%

despite 51% of facilities offering reuse channels [22]. The system possesses the technological capacity for high recovery but remains socially misaligned with the principles of a CE. In the Sri Lankan context, the resource recovery ecosystem operates primarily through informality and adaptability with limited formal infrastructure. In the absence of effective formal systems, a vast informal network has been operating nationwide. Table 4 summarises key characteristics of the UK and SL resource recovery systems. The UK and SL illustrate distinct circularity pathways, each characterised by partial success. The UK embodies top-down circularity built on policy frameworks, advanced technology, and infrastructure. SL represents bottom-up circularity, characterised by community-driven, local repair economies and informal material recovery.

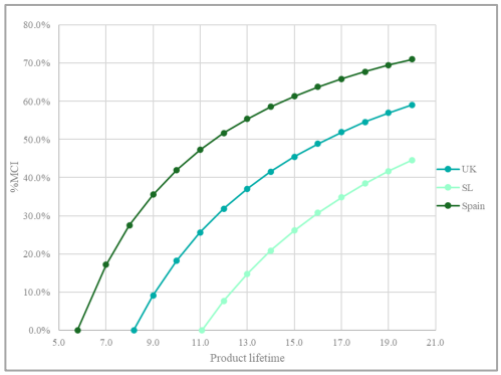


Figure 4: Relationship between product lifespan and MCI for washing machine

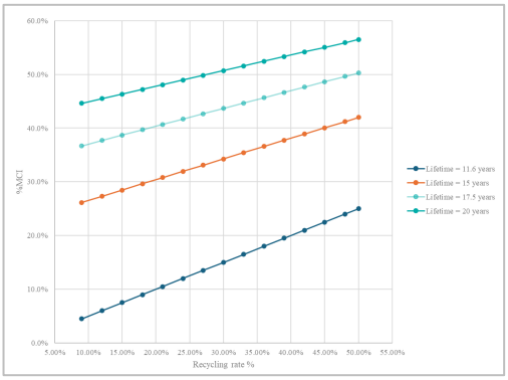


Figure 5: Impact of recycled content percentage on MCI across different product

Table 4: Comparative Analysis of UK and Sri Lanka Resource Recovery Systems

Dimension	United Kingdom	Sri Lanka
Reverse logistics	Extensive formal infrastructure, but poor collection behaviour	Decentralised informal door-to-door collection
Urban Mining	High-precision operations AI sorting, bioleaching, and advanced refining	Component-level manual disassembly No advanced sorting or domestic refining [24]
Domestic Refining Capacity	Operational: High-purity metal extraction	Absent: Most formal collections are exported for overseas processing [25]
System characterization	Top-down circularity	Bottom-up circularity
Strengths	Downstream processing efficiency Formal infrastructure	Upstream product preservation Extensive repair networks
Weaknesses	Upstream inefficiencies Weak reuse culture (2%) [22]	Downstream processing gaps Material recovery by acid leaching, open burning [26]
Health & Environmental Impact	Controlled	Hazardous: Informal sector uses unsafe methods (acid leaching, burning) [26]

6. Conclusion

Integration of UM, RL, and CLSC creates synergistic effects that substantially exceed the individual contributions to CE implementation. The proposed framework positions UM as the recovery mechanism, RL as the collection infrastructure, and CLSC as the system-level coordinator. This integrated decision-

support system guides EoL product management across multiple recovery routes and considers stakeholder interests at strategic, tactical, and operational levels. MCI quantifies the integrated approaches, achieving circularity scores 40%-50% higher than isolated implementations across product, component, and material recovery pathways. This advancement shows that integration among material identification, logistics, and supply-chain reintegration liberates value that is not available when these functions are operated separately.

The framework offers a hierarchical decision structure applicable to various situations, explaining where decisions are made, recovery-pathway criteria, and the responsibilities of stakeholders. This framework is designed for industrial decision-makers in manufacturing sectors who require integrated decision support for implementing the CE, integrating UM operations, RL planning, and CLSC management. The primary users include supply chain managers optimising recovery pathway selection and recovery facility operators assessing the feasibility of product, component, or material-level recovery strategies. The framework is operationalised through assessment of product condition and component value, RL network for collection, and technical constraints within the recovery pathway. Data on product returns, materials, recovery-facility capacities, and transportation networks would yield better results. The framework was tested on a washing machine to ensure it can be applied to other products that require sector-specific validation and adaptation. This extension would be heavily reliant on material data and recovery rates, which may be limited in some contexts where the informal sector is most dominant, particularly in developing economies. A carbon footprint analysis would further strengthen this work and should be considered in future research.

References

[1] Circle Economy Foundation, 2024, *The Circularity Gap Report 2024*. [Online]. Available: www.deloitte.com.

[2] Zengin, M., Amin, S. H., and Zhang, G., 2024, "Closing the Gap: A Comprehensive Review of the Literature on Closed-Loop Supply Chains," Logistics, 8(2). https://doi.org/10.3390/logistics8020054.

[3] Amir, S., Salehi, N., Roci, M., Sweet, S., and Rashid, A., 2023, "Towards Circular Economy: A Guiding Framework for Circular Supply Chain Implementation," Bus. Strategy Environ., 32(6), pp. 2684–2701. https://doi.org/10.1002/bse.3264.

[4] Biancolin, M., 2023, "Reverse Logistics and Circular Economy: A Literature Review," European Transport/Trasporti Europei, (94), pp. 1–14. https://doi.org/10.48295/ET.2023.94.7.

[5] Xavier, L. H., Ottoni, M., and Abreu, L. P. P., 2023, "A Comprehensive Review of Urban Mining and the Value Recovery from E-Waste Materials," Resour. Conserv. Recycl., 190. https://doi.org/10.1016/j.resconrec.2022.106840.

[6] Zeng, X., Mathews, J. A., and Li, J., 2018, "Urban Mining of E-Waste Is Becoming More Cost-Effective Than Virgin Mining," Environ. Sci. Technol., 52(8), pp. 4835–4841. https://doi.org/10.1021/acs.est.7b04909.

[7] Rogers, D. S., and Tibben-Lembke, R. S., 1999, *Going Backwards: Reverse Logistics Trends and Practices*, Pittsburgh, PA.

[8] Ellen MacArthur Foundation, 2025, *Keep It in Use: Retain Resource Value and Unlock Economic Opportunities*. [Online]. Available: https://www.ellenmacarthurfoundation.org/keep-it-in-use-retain-resource-value-and-unlock-economic-opportunities. [Accessed: 19-Dec-2025].

[9] Ottoni, M., Dias, P., and Xavier, L. H., 2020, "A Circular Approach to the E-Waste Valorization through Urban Mining in Rio de Janeiro, Brazil," J. Clean. Prod., 261. https://doi.org/10.1016/j.jclepro.2020.120990.

[10] Simaei, E., and Rahimifard, S., 2024, "AI-Based Decision Support System for Enhancing End-of-Life Value Recovery from e-Wastes," International Journal of Sustainable Engineering, 17(1), pp. 1–17. https://doi.org/10.1080/19397038.2024.2306293.

[11] Freitas, E. S. de C., and Xavier, L. H., 2025, "System Dynamics Applied to the E-Waste Value Chain: A Brazilian Case Study," Ecol. Modell., 506. https://doi.org/10.1016/j.ecolmodel.2025.111139.

[12] Tesfaye, F., Lindberg, D., Hamuyuni, J., Taskinen, P., and Hupa, L., 2017, “Improving Urban Mining Practices for Optimal Recovery of Resources from E-Waste,” Miner. Eng., 111, pp. 209–221. https://doi.org/10.1016/j.mineng.2017.06.018.

[13] Valai Ganesh, S., Suresh, V., Rajakarunakaran, S., and Suryakumar, V., 2025, “Sustainable Electronic Waste Management Framework for Academic Institutions in India,” Sci. Rep., 15(1), p. 40550. https://doi.org/10.1038/s41598-025-24278-z.

[14] Erdiaw-Kwasie, M. O., Abunyewah, M., and Baah, C., 2024, “A Systematic Review of the Factors – Barriers, Drivers, and Technologies – Affecting e-Waste Urban Mining: On the Circular Economy Future of Developing Countries,” J. Clean. Prod., 436. https://doi.org/10.1016/j.jclepro.2024.140645.

[15] Thinkstep-ANZ Services, “Thinkstep-ANZ. Material Circularity Indicator (MCI) Calculator.” [Online]. Available: https://www.thinkstep-anz.com/services/circular-economy/material-circularity-indicator-mci-calculator. [Accessed: 15-Dec-2025].

[16] Goddin, J., Marshall, K., Pereira, A., Design, G., Herrmann, S., Ds, S., Sam, J., Dupont, T., Krieger, C., Lenges, E., Ben, C., Pierce, J., Susan, E., Gispen, I.-J., Veenendaal, R., Per, I., Natureworks, S., Ford, L., Goodman, T., Vetere, M., Graichen, F., Anu, S., Tetrapak, N., and Cockburn, D., 2015, *Ellen MacArthur Foundation Project Team (in Alphabetical Order)*.

[17] Özkır, V., and Başlıgıl, H., 2012, “Modelling Product-Recovery Processes in Closed-Loop Supply-Chain Network Design,” Int. J. Prod. Res., 50(8), pp. 2218–2233. https://doi.org/10.1080/00207543.2011.575092.

[18] Boyano, A., Cordella, M., Espinosa, N., Villanueva, A., Graulich, K., Rüdenauer, I., F, A., and Stamminger, R., 2017, *Preparatory Study for Ecodesign and Energy Label for Household Washing Machines and Washer Dryers*. https://doi.org/10.2760/029939,.

[19] DTP Group, “50+ E-Waste Statistics 2025.” [Online]. Available: https://dtpgroup.co.uk/insight/50-e-waste-statistics-2025/. [Accessed: 18-Dec-2025].

[20] “Ceylon Waste Management.” [Online]. Available: https://ewaste.lk/sustainability.php. [Accessed: 18-Dec-2025].

[21] Talens Peiró, L., García Fernández, B., and Gabarrell i Durany, X., 2022, “Investigating a Repair Workshop: The Reuse of Washing Machines in Barcelona,” Sustain. Prod. Consum., 29, pp. 171–179. https://doi.org/10.1016/j.spc.2021.10.003.

[22] The Restart Project, 2024, “Reuse Streams at UK Waste Facilities (HWRCs).” https://doi.org/10.5281/zenodo.13930139.

[23] Dyson, G., “How Long Do Washing Machines Last? Homebuilding.” [Online]. Available: https://www.homebuilding.co.uk/rooms/utility/how-long-do-washing-machines-last. [Accessed: 18-Dec-2025].

[24] Gunarathne, N., de Alwis, A., and Alahakoon, Y., 2020, “Challenges Facing Sustainable Urban Mining in the E-Waste Recycling Industry in Sri Lanka,” J. Clean. Prod., 251. https://doi.org/10.1016/j.jclepro.2019.119641.

[25] Ranasinghe, W. W., and Athapattu, B. C. L., 2019, “Challenges in E-Waste Management in Sri Lanka,” *Handbook of Electronic Waste Management: International Best Practices and Case Studies*, Elsevier, pp. 283–322. https://doi.org/10.1016/B978-0-12-817030-4.00011-5.

[26] Widanapathirana, S., Perera, I. J. J. U. N., and Bellanthudawa, B. K. A., 2023, “Electrical and Electronic Waste (e-Waste) Recycling and Management Strategies in South Asian Region: A Systematic Review from Sri Lankan Context,” Waste Dispos. Sustain. Energy, 5(4), pp. 559–575. https://doi.org/10.1007/s42768-023-00162-2.

Biography

Themiya Kuruppuge is a PhD candidate at the University of Exeter, with extensive experience in environmental services and process engineering. His current research focuses on advancing sustainable manufacturing practices to support the transition toward a circular economy in the digital era. His e-mail address is tk504@exeter.ac.uk

Asela K. Kulatunga is a Lecturer at the Department of Engineering, University of Exeter, UK. He holds a Ph.D. in Industrial Engineering from the University of Technology, Sydney, Australia. He currently serves as the Director of the MSc Programme in International Supply Chain Management. His research interests are the sustainability aspects of industrial and supply chain engineering problems. His e-mail address is a.k.kulatunga@exeter.ac.uk

Martino Luis is a Senior Lecturer at the Department of Engineering, University of Exeter, UK. He holds a Ph.D. in Operations Research from Kent Business School, University of Kent, UK, and co-leads the ExDES laboratory with Professor Sucala. His research interests are the applications of optimisation and simulation modelling in facility location, supply chain design, and inventory routing with a focus on sustainable operations. His e-mail address is m.luis@exeter.ac.uk

Voicu Ion Sucala is a Professor in Engineering Management and the Head of Engineering at the University of Exeter. He holds a Ph.D. in Industrial Engineering (TU Cluj-Napoca) and a Ph.D. in Social Sciences (University of Glasgow), serves as Chair of the Research, Innovation and Knowledge Transfer committee at the Engineering Professors' Council, and co-leads the Exeter Digital Enterprise Systems (ExDES) laboratory. His research focuses on modelling, simulation, digital twinning, and optimisation. His email address is i.sucala@exeter.ac.uk

8th Conference on Production Systems and Logistics

Enabling Human-AI Collaboration in a Modular Planning Environment

Esben Schukat[1], Lea Röhm[1], Markus Netzer[2,3], Günther Schuh[1]

[1] *Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Germany*
[2] *NG GmbH, Amtzell, Germany*
[3] *Aalen University of Applied Science, Germany*

Abstract

The vision of fully automated factory planning remains an elusive goal. The complexity and context-dependency of tasks in factory panning still demand human expertise. This paper introduces the concept of a Factory Planning Copilot, which shifts the focus from automation to intelligent assistance. Inspired by systems like Microsoft Copilot, the approach envisions a virtual environment where planners interact with digital factory models through natural language, supported by large language models. Rather than replacing human judgment, the Copilot augments it by interpreting intentions, suggesting improvements and enabling intuitive access to complex planning data. By embedding domain knowledge through modular expert systems and fostering an open ecosystem, the Copilot can accelerate planning processes, enhance decision quality, and improve compliance with internal and external requirements. Copilot for Factory Planning aims to make digital factory planning more accessible, adaptive, and reliable to empower planners through seamless human-AI collaboration.

Keywords

Factory Planning; Artificial Intelligence; Human-AI Collaboration; Copilot; Digital Factory Planning

1. Introduction

Global geopolitical shifts and economic challenges have heightened the need for resilient and efficient production infrastructures. Crises such as the COVID-19 pandemic exposed vulnerabilities in global supply chains and accelerated structural changes in production networks, prompting a move towards re- and nearshoring strategies. [1-3] These developments not only relocate production but also increase the volatility of conditions under which factories are planned and operated, transforming factory planning into a highly dynamic process. Ambitious climate targets further require the rapid deployment and scaling of new technologies, creating a dense network of interdependencies. Planning must integrate technological, economic, and ecological criteria while remaining flexible to disruptions or regulatory changes. [4]These challenges are compounded by a severe shortage of skilled labor. In 2022, Germany reported 630,000 unfilled positions in factory planning and construction, causing an estimated €100 billion loss in value creation. By 2030, the retirement of millions of experienced professionals will exacerbate this scarcity, creating bottlenecks as task complexity increasingly exceeds human capacity.[5-7]

Factory planning thus becomes a central lever for addressing these challenges, as it shapes how infrastructures can remain both robust and adaptive under dynamic conditions. [8] Although digital tools already improve accuracy and collaboration, their integration has so far been limited. Fundamental issues

DOI: https://doi.org/10.15488/21020
ISSN: 2701-6277

such as fragmented information flows, insufficient coordination, and high complexity persist, meaning that technological progress has not yet translated into a substantial transformation of factory planning.[9]

These shortcomings point to the limits of incremental improvement and underline the need for more transformative approaches. Artificial intelligence offers such potential by acting not merely as another isolated tool but as an assistance system that integrates fragmented processes, interprets complex data, and supports decision-making in real time. Conceptualized as a Factory Planning Copilot, AI-driven systems could complement human expertise, highlight interdependencies, and provide contextualized guidance, thereby enabling a more adaptive, accessible, and resilient planning process.[10-12]

2. Factory Planning: A look into practice

2.1 Digital tools for Factory Planning

Factory planning is a structured process that encompasses the design and organization of production systems. It includes tasks such as facility layout, workplace arrangement, material flow optimization, and equipment selection. [13, 14] These activities are distributed across successive phases, beginning with conceptual planning, continuing with detailed planning, followed by implementation, and extending into the operational phase. Each phase involves distinct challenges and decision-making requirements, which must be coordinated across multiple disciplines and stakeholders to ensure efficient and effective production processes.[13]

Digitalization provides the technological foundation to support these complex planning activities. It enables planners to structure information, model production systems, and evaluate alternative solutions. Digital approaches promise increased transparency, efficiency, and flexibility, offering tools that can assist planners at different stages of the planning process. [15] Computer-Aided Design applications support the creation of technical drawings and spatial visualizations, facilitating the evaluation of alternative layouts during conceptual planning and the precise specification of workplaces and equipment in detailed planning. [15] Building Information Modelling represents a methodological framework rather than a single software solution. It integrates geometric and semantic information across the facility lifecycle, fosters interdisciplinary collaboration, ensures documentation consistency, and reduces redundancies. Its methodological character is particularly valuable in the transition from detailed planning to implementation and operation, where multiple disciplines and stakeholders must align their models and information. [16] Simulation software enables the virtual exploration of dynamic processes. During conceptual planning, it allows the comparison of alternative system configurations and assessment of throughput scenarios. In later phases, simulation supports the validation of detailed designs, the optimization of workflows, and the evaluation of planned changes prior to implementation. By providing decision support across all phases, simulation contributes to risk reduction and enhances flexibility in factory transformation. [17]

Taken together, CAD, BIM, and simulation constitute a mature and powerful digital toolbox for factory planning. In practice, however, these instruments are often applied selectively and in isolation. Instead of forming an integrated planning environment, they typically exist as separate solutions with limited interoperability. This fragmentation generates information gaps, inefficiencies, and difficulties in translating digital models into actionable planning support, highlighting the need for approaches that integrate these tools more systematically.

2.2 Challenges and Limitation of Digital Factory

Survey data indicate that despite their potential, digital tools remain underutilized in factory planning. In 2023, fewer than 60% of companies used 2D or 3D CAD software, and under 45% employed advanced applications such as material flow simulation or layout optimization, highlighting a persistent gap between availability and adoption. Historical comparisons show that core challenges, particularly the lack of

synchronization between planning tasks, causing information fragmentation and time losses, have changed little over the past decade. Limited personnel qualifications, scarce planning capacities, and an unclear cost-benefit ratio remain major obstacles, with over 70% of companies reporting skill-related barriers and more than 60% citing financial concerns. [18, 19]

Beyond these structural issues, current tools primarily generate and manage data but provide little support in interpreting it within the broader decision context. Planners still need to integrate heterogeneous information, evaluate trade-offs, and make decisions under uncertainty, so increased data availability does not automatically improve outcomes. Incremental improvements of existing tools are insufficient. What is required are systems that actively assist in making sense and connect different forms of data. Such systems would complement, not replace, established tools like CAD, BIM, or simulation by enabling intuitive access to complex data, suggesting context-sensitive improvements, and coordinating information flows across applications.[19, 20]

This perspective motivates the exploration of AI-driven assistance systems as a transformative step in digital factory planning. By leveraging large language models and modular expert systems, AI can serve as an integrative layer between fragmented tools and human expertise, empowering planners by interpreting intentions, highlighting interdependencies, and making planning knowledge more actionable.

3. Artificial Intelligence in Digital Factory Planning

3.1 Definition of Artificial Intelligence

The limitations of current digital planning tools highlight the need for systems that go beyond mere data provision and fragmented workflows. Planners increasingly face complex interdependencies that demand not only access to information but also context-aware guidance to support efficient and informed decision-making.

Artificial intelligence encompasses a range of technologies designed to emulate human cognitive functions, including reasoning, learning, and problem-solving.[8] Its core domains serve specific purposes and have already been successfully applied in multiple sectors. Text and speech processing enables intuitive human–machine interaction by interpreting natural language and semantics, as exemplified by tools such as eDiscovery in the legal sector, which automatically structures and categorizes large volumes of documents to support case preparation and analysis. Knowledge representation assigns meaning to data and models logical relationships, allowing the interpretation of large and complex datasets, as employed in the financial sector for automated fraud detection and risk assessment. Video and image processing facilitates the analysis of visual information, which is widely used in healthcare for diagnostic imaging and in manufacturing for quality inspection. [8] Pattern recognition uncovers hidden structures and correlations within data, while action planning and optimization supports decision-making in complex environments. Emotion recognition allows systems to interpret human affective states through text, speech, or facial expressions. Collectively, these capabilities illustrate the broad potential of AI to replicate and augment human-like thinking and decision-making processes, providing a foundation for more adaptive, intelligent, and context-aware planning tools. While AI has already proven effective in domains such as law, finance, healthcare, and manufacturing, its adoption and systematic integration within factory planning remain the subject of ongoing research, highlighting both opportunities and challenges for future applications.[21]

3.2 Research on Artificial Intelligence in digital Factory Planning

While AI has demonstrated its problem-solving capabilities across various industries, its application in factory planning remains at an early stage. Initial integration in the 1990s involved rule-based expert systems that supported specific planning tasks. [22, 23] With the advancement of technologies and the growing availability of data, machine learning and neural networks were subsequently employed to address more

complex challenges. [23] Today, AI provides the ability to dynamically adapt to rapidly changing planning conditions that increasingly exceed human capacities.[23, 24]

A systematic literature review conducted in November 2023 across Scopus, IEEE Xplore, and ScienceDirect confirms this limited adoption. From an initial set of 609 results, 34 articles were selected for full-text review, of which only seven demonstrated concrete applications of AI to factory layout or design. Two studies investigated the processing of near-real-time production data within digital factory twins [25, 26], four addressed layout and capacity planning [27-30], and one explored the conversion of three-dimensional scans into CAD models [31].

These findings indicate that AI in factory planning has largely focused on isolated technologies and specific subtasks, while integrated solutions and real-world implementations remain scarce. This gap highlights both the relative immaturity of the field and the opportunity for transformative approaches. Systems that combine the analytical capabilities of AI with intuitive, planner-centered interaction, have the potential to move beyond incremental improvements, enabling planners to manage complexity, identify critical interdependencies, and render digital planning more adaptive, accessible, and resilient.

4. Artificial Intelligence in Digital Factory

4.1 Concept: Factory Planning Copilot

Building on the challenges outlined in the previous chapters, the Factory Planning Copilot is introduced as an intelligent, planner-centered system that unifies analytical capabilities, domain-specific expert modules and large language models (LLMs) within a single operational environment. This integration enables planners to investigate, modify and optimize production scenarios through natural language interaction. The Copilot acts as an adaptive orchestration layer that interprets planning intentions and translates them into targeted computational procedures, thereby reducing the need for direct engagement with specialized software and supporting the execution of complex planning operations through intuitive dialogue. The overall conceptual structure of the copilot is presented in Figure 1. By embedding domain knowledge in modular expert systems and by adopting an open, extensible design, the Copilot supports context-aware guidance and iterative feedback, enhancing decision quality while ensuring adherence to internal guidelines and external regulatory frameworks. The extensible knowledge base allows the system to incorporate additional expert systems, analytical procedures or rule sets as new planning requirements arise, ensuring long-term adaptability.

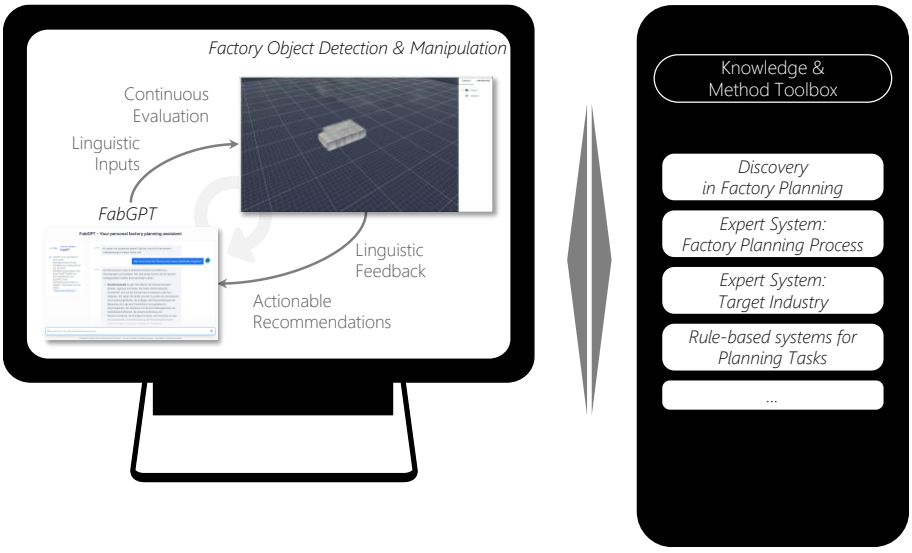


Figure 1: Conceptual schematic of the copilot's high-level structure.

The system integrates a range of established planning tools, including simulation environments, CAD and BIM platforms and diverse optimization modules. When planners articulate questions, planning intents or modification requests, the Copilot evaluates these inputs and coordinates the associated actions within the modular ecosystem. This process begins with the LLM interpreting the natural-language input and identifying the relevant domain knowledge within the Knowledge and Method Toolbox. The LLM then initiates the corresponding layout adaptations, analyses or simulations in the connected technical systems, thereby coupling linguistic interpretation with computational execution. Through this mechanism, factory planning becomes a continuous dialogue in which human expertise is augmented by AI-driven reasoning. The bidirectional exchange between the LLM and the simulation or layout engines ensures that technical outputs, such as geometric modifications, performance metrics or compliance findings, are returned to the LLM. These outputs are subsequently contextualized and synthesized into coherent recommendations that can be directly applied to the planning task.

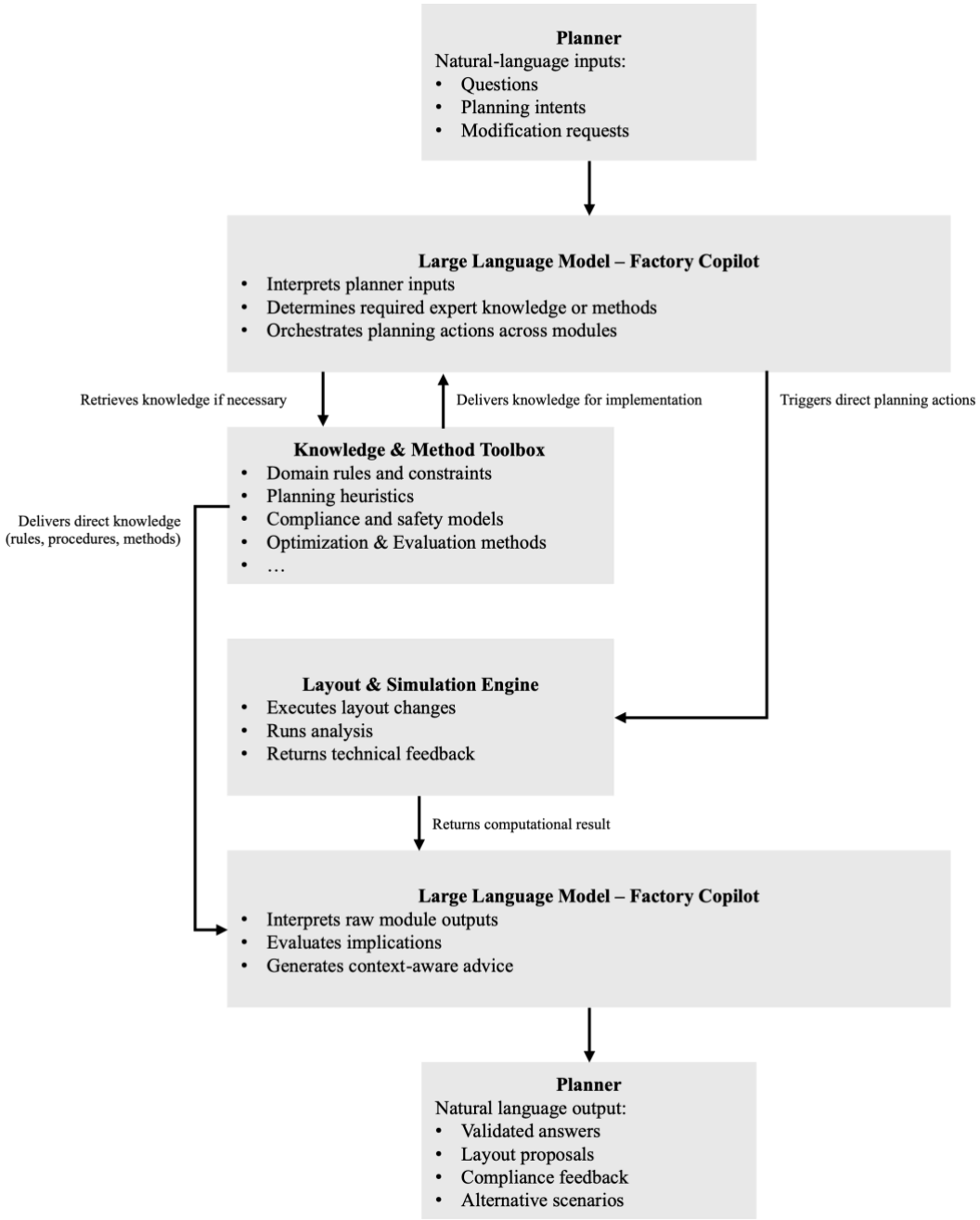


Figure 2: Interaction of the Factory Planning Copilot with its modular open ecosystem.

Figure 2 illustrates the modular ecosystem of the Factory Planning Copilot and the interaction between planner inputs, the LLM layer and the embedded expert modules. The LLM receives the planner's natural language requests and determines which domain knowledge or methodological procedures are required to address the planning task. In addition to initiating direct modifications within the layout or simulation environment, the Copilot also enables planners to access the Knowledge and Method Toolbox solely for the purpose of retrieving expert knowledge, such as standards, regulatory information or methodological guidance, without triggering operational changes in the planning model. This knowledge is obtained from the open and extensible Knowledge and Method Toolbox, which provides domain rules, analytical procedures and expert methods. The Copilot integrates this information into its reasoning process and subsequently translates it into structured commands that are executed within the connected CAD, BIM or simulation systems. These systems return computational results, such as geometric modifications, performance indicators or compliance-related findings, to the LLM layer, where they are evaluated and synthesized into coherent, context-sensitive recommendations for the planner.

The open design of the Knowledge and Method Toolbox ensures that new analytical components or domain-specific expert systems can be incorporated without altering the overarching architecture, thereby supporting continuous methodological expansion. Through the consolidation of expert knowledge, simulation outputs and planning intent, the LLM generates iterative feedback that allows planners to refine and validate their decisions in a continuous, dialogue-based workflow. This iterative process establishes a planner-centered interaction pattern in which the Copilot progressively adapts its reasoning to updated model states, to newly integrated expert modules and to evolving planning objectives, ensuring that its guidance remains both responsive and grounded in the available domain knowledge.

4.2 Large Language Models for Human-Machine Interaction

Large Language Models form the core of the Copilot's interaction layer. They allow planners to communicate assumptions, requirements, and queries naturally, without the need for specialized interfaces. LLMs interpret planner input, contextualize heterogeneous data, reveal interdependencies, and highlight potential conflicts or optimization opportunities.

Beyond translating instructions into system-specific operations, LLMs provide real-time validation of assumptions, suggest alternative approaches, and generate actionable guidance. This enables an iterative, dialogue-driven planning process in which human expertise remains central while benefiting from computational insights. By connecting fragmented tools and modules, LLMs facilitate a seamless, planner-centered interaction across the entire planning ecosystem.

4.3 Knowledge and Method Toolbox

The Copilot's Knowledge and Method Toolbox consists of specialized expert modules that incorporate domain-specific knowledge, empirical data, regulatory standards, and optimization algorithms. These modules can be dynamically integrated depending on the planner's objectives, providing scenario-specific support. For instance, a layout optimizer can be activated when space utilization becomes critical, while a sustainability module can guide energy-efficient configurations.

The LLM consolidates outputs from all modules, presenting results as coherent, context-aware recommendations. This approach supports iterative assumption testing, validation, and continuous refinement, enabling planners to formulate hypotheses regarding production layouts, resource allocation, or process changes, receive immediate feedback, and explore alternative solutions efficiently while maintaining full control over decision-making.

4.4 Use Case: Adaptive, Interactive Planning

A practical scenario illustrates the Copilot's functionality. Consider a planner who must verify whether the current layout of a production area complies with safety regulations, such as the escape route requirements

defined in the workplace ordinance. Instead of manually inspecting distances or consulting several independent tools, the planner formulates the requirement in natural language by requesting an evaluation of whether all escape routes comply with the relevant regulatory standards and by asking for feedback on potential improvements.

The Copilot interprets this instruction and identifies which regulatory information, methodological procedures and analytical steps are necessary to conduct the assessment. It accesses the Knowledge and Method Toolbox to retrieve the applicable safety standards and the corresponding evaluation methods and subsequently orchestrates the relevant compliance-checking and spatial-analysis modules. The connected simulation or layout environment performs the required computations and returns the resulting indicators, geometric analyses and compliance findings to the LLM layer. The Copilot integrates this operational output with the previously retrieved regulatory knowledge and synthesizes it into a coherent and context-sensitive assessment that clearly communicates compliant and non-compliant areas as well as opportunities for improvement.

If the planner introduces additional considerations, such as the need to optimize space utilization or reduce internal travel distances, the Copilot dynamically incorporates further expert modules by drawing on the extensible Toolbox and invoking the appropriate analytical or optimization routines. This capability enables the planner to explore alternative solutions, examine trade-offs and iteratively refine the layout based on updated analytical insights. Throughout this process, decision-making remains firmly in the hands of the planner, while the Copilot provides integrated analyses, domain knowledge and continuous feedback that support transparent justification and informed planning decisions.

5. Conclusion and Outlook

This paper has presented the Factory Planning Copilot as a novel, planner-centered approach for interactive, dialogue-driven factory planning. By placing natural language communication at the core of human–system collaboration, planners can formulate goals, generate and validate assumptions, and receive context-aware guidance. By embedding domain knowledge through modular expert systems and fostering an open ecosystem of methods and tools, the Copilot consolidates fragmented planning tools into a coherent assistance environment, enhances decision quality, and supports compliance with internal and external requirements.

The Copilot transforms planning from a sequence of isolated tasks into an iterative, collaborative process in which assumptions can be tested, alternatives explored, and interdependencies revealed in real time. Its modular architecture ensures flexibility and scalability, enabling planners to selectively integrate domain-specific methods, expert systems, or optimization algorithms while keeping human judgment central. Transparent reasoning, context-sensitive interpretation, and high-quality data remain critical prerequisites for trust and practical adoption.

A demonstrator is being developed using Unreal Engine to enable interactive scenario exploration. Defining platform rules will allow planners to dynamically integrate external methods and tools. This open, extensible ecosystem of methods and expert modules lays the foundation for adaptive, knowledge-rich, and resilient planning. Future work will focus on refining contextual interpretation, expanding dynamic module integration, and validating the approach in real-world environments.

References

[1] ABB, 2023. ABB nennt die wichtigsten Robotik-Trends für 2023.

[2] Kessler, S., 2021. US-Firmen holen Produktion zurück, in: Deutsche Welle.

[3] Dullien, S., 2021. Nach Corona kommt die Deglobalisierung.

[4] Umweltbundesamt, 2023. Projektionsbericht 2023 für Deutschland.

[5] Miegel, B., 2022.ABB-Studie zeigt: weltweite Ausbildungslücke gefährdet nächsten Industrialisierungsschub.

[6] Tiedemann, J., Malin, L., 2023. Jahresrückblick 2022 - Fachkräftesituation angespannter denn je: Studie im Rahmen des Projektes Kompetenzzentrum Fachkräftesicherung (KOFA) in Zusammenarbeit mit dem Bundesministerium für Wirtschaft und Klimaschutz (BMWK).

[7] Statista, 2020. Bis 2030 fehlen in Deutschland 2 Millionen Fachkräfte.

[8] Hatiboglu, B., Schuler, S., Bildstein, A., Hämmerle, M., 2019. Einsatzfelder von künstlicher Intelligenz im Produktionsumfeld, in: Fraunhofer Institute for Manufacturing Engineering and Automation IPA.

[9] VDI Expert Forum: Survey within the Expert Forum on Factory Planning on VDI 5200, 2023.

[10] Dell' Acqua, F., McFowland III, E., Mollick, E., Lifshtz-Assaf, H., Kellogg, K.C., Rarendran, S., Krayer, L., Candelon, F., Lakhani, K.R., 2023. Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality, in: Harvard Business School Technology & Operations Mgt. Unit Working Paper, vol. No. 24–013.

[11] Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Gorecky, D., 2016. Towards an Operator 4.0 Typology: A Human-Centric Perspective on the Fourth Industrial Revolution Technologies, in: International Conference on Computers & Industrial Engineering (CIE46).

[12] Link, M., Hamann, K., 2019. Einsatz digitaler Assistenzsysteme in der Produktion, in: ZWF, vol. 114(10), pp. 683-687.

[13] Burggräf, P., Schuh, G., 2021. Fabrikplanung, in: Springer Berlin Heidelberg.

[14] Pawellek, G., 2024. Ganzheitliche Fabrikplanung.

[15] Dannapfel, M., 2023. Factory Planning Lecture at RWTH Aachen University.

[16] Latiffi, A. A., Brahim, J., Fathi, M. S., 2014. The Development of Building Information Modeling (BIM) Definition, in: Applied Mechanics and Materials, vol. 567, pp. 625-630.

[17] Burggräf, P., Adlon, T., Beyer, M., Xu, T. a., 2022. Digitale Fabrikplanung als Befähiger zur Erreichung der Klimaziele/Scalable factory and production concept for the production of water electrolysers, in: wt Werkstattstechnik online, vol. 112, no. 04, pp. 210-215.

[18] Center Factory Planning: Working Group on Factory Planning, 2023.

[19] Neuhäuser, T., Spiegelsperger, S. M., Hohmann, A., Daub, R., Weist, K. C., 2023. Relevanz von BIM in der Fabrikplanung/Building information modeling in factory planning -- Results of an online survey, in: wt Werkstattstechnik online, vol. 113, no. 03, pp. 93-100.

[20] Reinema, C., Pompe, A., Nyhuis, P., 2013. Agiles Projektmanagement, in: ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, vol. 108, no. 3, pp. 113-117.

[21] Kreutzer, R. T., Sirrenberg, M., 2019. Künstliche Intelligenz verstehen: Grundlagen-Use-Cases - unternehmenseigene KI-Journey.

[22] Gao, R. X., Krüger, J., Merklein, M., Möhring, H.-C., Váncza, J., 2024. Artificial Intelligence in manufacturing: State of the art, perspectives, and future directions, in: CIRP Annals, vol. 73, no. 2, pp. 723-749.

[23] Rashid, A. B., Kausik, M. A. K., 2024. AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications, in: Hybrid Advances, vol. 7, p. 100-277.

[24] Eschemann, P., Oppenheimer, F., 2022. Effiziente Fabriken dank künstlicher Intelligenz, in: ti-Magazin.

[25] Kuehn, W., 2018. Digital twins for decision making in complex production and logistic enterprises, in: International Journal of Design, Nature and Ecodynamics, vol. 13, no. 3, pp. 260-271.

[26] Wang, H., Du, W., Li, S., 2022. Key Issues for Digital Factory Designing and Planning: A Survey, in:. Software Engineering, Artificial Intelligence, vol. 1012, pp. 18-29.

[27] Burggräf, P., Wagner, J., Heinbach, B., 2021. Bibliometric Study on the Use of Machine Learning as Resolution Technique for Facility Layout Problems, in: IEEE Access, vol. 9, pp. 22569-22586.

[28] Ceyhan, e., Zkan, Y., Ener, E., Etin, G. Z.,2023. A Flexible Facility Layout Design for a Project-Based Production System, in: Towards Industry 5.0, pp. 686-696.

[29] Knollmeyer, S., 2023. Ontology based knowledge graph for information and knowledge management in factory planning.

[30] Pinto, R., Goncalves, J., Cardoso, H. L., Oliveira, E. G., 2016. A Facility Layout Planner tool based on Genetic Algorithms.

[31] Sommer, M., 2023. Automated generation of digital twin for a built environment using scan and object detection as input for production planning, in: Journal of Industrial Information Integration, vol. 33, p. 100462

Biography

Dr. -Ing. Esben Schukat (*1993) has been Chief Engineer in Department for Factory Planning at the Chair of Production Engineering at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2023. He is also Managing Director of the Center Factory Planning and the Center for Smart Industrial Agriculture as well as Senior Manager of the Center Smart Assembly.

Lea Röhm (*1998) has been student research assistant at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2021.

Dr. -Ing. Markus Netzer (*1993) earned his doctorate at the wbk Institute of Production Science at the Karlsruhe Institute of Technology (KIT), where he specialized in artificial intelligence in production. He is a lecturer at Aalen University of Applied Sciences, teaching production management. In addition, he serves as managing director of a technology-oriented company focused on transferring innovative machining technologies into industrial practice.

Prof. Dr. -Ing. Günther Schuh (*1958) has been head of the Chair of Production Engineering at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2002. He is Director of FIR e.V. as well as a Member of the Board of WZL and Fraunhofer IPT. He is also a member of numerous technical-scientific committees and founder and CEO of several start-ups in the field of electro mobility.

8th Conference on Production Systems and Logistics

A Chinese Postman Problem-Based Approach For Municipal Solid Waste Management

Akshat Kothyari[1], Shaunak Juvekar[2], Subhash C. Sarin[1]
[1]*Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, USA*
[2]*Department of Computer Science, Virginia Tech, Blacksburg, USA*

Abstract

In this paper, we address a municipal solid waste (MSW) collection problem and formulate it as a multi-echelon network design problem inspired by the Chinese Postman Problem (CPP). The waste is collected along the streets, and the objective is to achieve a complete street-level coverage for waste collection while minimizing the total system cost. The proposed formulation captures the hierarchical structure of the MSW logistics that spans collection routes, transfer stations, and location of processing facilities. To address the combinatorial complexity of the problem, a column-generation-based decomposition procedure is developed, where the master problem involves selection of the routes that cover mutually exclusive sets of streets via mutually exclusive sets of facilities at a higher echelon. The subproblem iteratively determines better vehicle routes. The approach ensures computational scalability for realistic urban networks with thousands of street segments. A web-based implementation of the proposed methodology integrates ESRI routing APIs to retrieve network distances and travel times, thereby enabling geographically accurate and adaptable analyses for different urban layouts. This platform allows users to evaluate design scenarios under varying waste generation patterns and vehicle and facility capacities, thereby, supporting the development of circular and resource-efficient waste management systems. By combining exact optimization techniques with spatial analytics, we effectively demonstrate their applicability in sustainable urban infrastructure planning, illustrating how classical routing theory can be extended to modern challenges in circular economy systems.

Keywords

Mathematical programming; Municipal solid waste management; Column generation; Circular economy.

1. Introduction

The increasing challenge of global solid waste management, driven by rapid population growth and urbanization, necessitates a critical shift toward sustainable and efficient systems that align with circular economy principles. Traditional waste disposal models are economically inefficient, with collection and transportation often accounting for the majority of operational costs. Optimizing these processes is a key strategy for reducing fuel consumption, emissions, and overall expenditures, thereby enhancing urban sustainability. This paper addresses the design of an optimized solid waste management network that facilitates circular economy, focusing specifically on the vehicle routing problem (VRP) inherent in the collection phase. Utilizing a model formulation based on the well-known Chinese Postman Problem (CPP), which seeks the shortest possible route to traverse all required street segments in a network, this research proposes a method to significantly improve operational efficiency and ensure a comprehensive service coverage. The proposed methodology is then incorporated in a web-based tool that incorporates GIS information and produces optimally designed MSW collection network. This provides a robust framework

DOI: https://doi.org/10.15488/21021
ISSN: 2701-6277

for municipalities to transition from historical, inefficient routing practices to data-driven, cost-effective, and environmentally sound waste recovery networks.

2. Literature Review

The circular economy is gaining recognition as a viable pathway toward sustainable development, emphasizing waste management strategies that favour reduction, followed by reuse, recycling, energy recovery, and ultimately, disposal [16]. Early studies in MSW management emphasized cost-effective collection and disposal, primarily modelling waste flows through landfills or incinerators. Over time, the research agenda evolved to incorporate environmental criteria, resource recovery, and the allocation of waste across multiple processing and treatment facilities [2]. This shift aligns with global movements toward sustainable development and the circular economy, where a system-wide optimization supports material recirculation, reduced landfill dependence, improved resource efficiency, and other economic benefits [14].

MSW management Activities: With the considerations of sustainability, resource recovery and segregation of the waste, the activities encountered in a MSW system include generation, storage, collection, transfer and transport, processing, and disposal of solid waste ([19], [27], and [15]). The sources of MSW include residential, commercial, institutional, construction and demolition, and urban services [8]. In many countries, most of the MSW is dumped in landfills ([29]), which serve as disposal sites for the system. Processing of waste may involve activities like recycling, conversion of waste to energy ([24]) or another treatment alternative, at dedicated treatment centers. To transfer waste from generation areas to treatment centers, intermediate transfer stations are often used in a MSW system ([30]). At the transfer stations, the waste is possibly compressed, and then hauled to the treatment centers or disposal sites using trailers. Similar multi-level facility-based systems are observed in hazardous waste management systems ([3]).

MSW management network design: The components of a MSW network can be broadly categorized as: (1) sources of waste, (2) facilities which represent waste treatment/processing facilities, transfer stations, or landfills, and (3) collection vehicles which transfer waste between (1) and (2). The cost is incurred in a MSW system because of routing of collection vehicles, location of facilities, and loading/waiting time ([13]). The transfer/collection decisions involve allocating the generation points to the facilities optimally. The collection from generation points is carried out by vehicles that transport the collected waste to the facilities. In most of the studies reported in the literature, these decisions are carried out by either allocating the generation points directly to the facilities ([12]) or by determining vehicle routes that traverse through generation points and deliver the waste to the facilities ([25]). In the presence of a hierarchy, the interactions between each level of facilities can be accomplished either via direct allocation or via routing ([7]). In [26], the author presents a multi-level industrial hazardous waste management problem, involving routing decisions between generation points, transfer stations, recycling centers and disposal centers.

MSW management network formulations: In the recent literature, the MSW problem has been formulated as a Location Routing Problem (LRP). A LRP is a location problem where the service to customers is provided by a fleet of vehicles in less-than-truckload routes ([17]). In [22], the authors have presented a multi-objective industrial hazardous waste LRP and developed an evolutionary-algorithm-based heuristic for its solution. Previously, [31] presented an instance of industrial/hazardous LRP with an added objective of minimizing the travel risk. On the other hand, the LRP-based MSW systems applied to a real-world scenario are presented in [5], [4] and [6]. Although all these models are LRP based, they differ in model formulation. In terms of the coverage of collection, some case studies have viewed the collection on a street-wise basis instead of on a point-wise basis. This is pertinent to cases where the number of sources of waste in a street is large. The street-wise collection problem can be viewed as the CPP, which has been studied in the past for designing MSW management networks ([10], [28], and [20]). As noted earlier, the MSW systems applied to real-life situations exhibit a multi-level structure. A variant of the LRP that has a similar multi-level structure, is called a Multi-Echelon LRP (ME-LRP) where the flow of material occurs between

different levels. Other common terms for this problem found in the literature are multi-stage, multi-layer, multi-level, and multi-stage LRP ([21]).

2.1 Contributions of this work

All the works presented above focus on a particular application, and therefore, need substantial changes in model formulation for implementation in other environments. Also, the discrete set of available capacities for facilities and consideration of multiple echelons are not addressed in one model formulation. The novelty of our work is in its comprehensibility thereby enabling its use to simultaneously: (i) design a MSW network of variable number of echelons, (ii) locate facilities of various sizes, (iii) effectively route the vehicles for collection of wastes, (iv) handle different types of wastes, and (v) use a route-based API for determination of travel time and distances. We also present a new column-generation-based method to solve this comprehensive model formulation.

3. Problem Description

In this section, we formally present the MSW problem that we address in this paper. As noted earlier, we consider a Multi-Echelon Location Routing Problem (ME-LRP) type structure for this system. For a single depot level ME-LRP, the key decisions involve locating the depots and deciding Hamiltonian paths for different vehicles that originate from these depots, traverse target streets to collect the material, and then route the vehicles back to their respective depots. The same set of decisions are involved when we include a child depot node level in between the parent depot level and target node level for a multi-level ME-LRP but instead of street-wise coverage, it follows a node-wise coverage since collections have to be made from lower echelon's depot level facilities. Hence, each level involves a similar set of decisions, and thereby, a similar set of components is required to solve this problem that differ by their hierarchy in the system. These components constitute different sets of locations, routing outcomes and costs related to their realizations. In this paper, the ME-LRP caters to collection of different types of products. The selected depots have the ability to hold all of these product types, and the collection strategies for the products are independent.

To that end, we first provide a mathematical model formulation for the ME-LRP encountered in managing the MSW and designate it as ME-LRP(MSW). The notation that we use can be found in Appendix A1. The cost incurred in locating a facility is due to purchasing of the land, or building the facility, and therefore, can be treated as a one-time investment in setting up a multi-echelon collection network. On the other hand, the cost incurred for routing the vehicles to collect the material and the cost incurred for operating the facility that handles this material, are recurring costs and are incurred whenever the collection is being performed. For simultaneous consideration of the non-recurring and recurring costs, an appropriate weight should be assigned to capture the total cost incurred in establishing the network for a given time horizon. If the routing considerations differ between different time periods (implying varying operation costs over time periods), then an aggregation of the routing and operation costs incurred constitutes the total cost of setting up the ME-LRP network for a given time horizon.

In this paper, we focus our attention on the ME-LRPs that incur only the recurring costs. This type of formulation is pertinent for systems that serve the same set of targets over different time periods. A system of this type is observed in Municipal Solid Waste (MSW) management where the waste is collected from the same set of target nodes periodically. Figure 1 illustrates the framework for a MSW management system where the collection occurs across three echelons (depot levels). In the figure, these depot levels are called, namely, Bin Locations, Site Locations and Central Hub in an increasing order of hierarchy. From a practical viewpoint, the Central Hub can be a waste processing center or a landfill. In this system, $T = 3$. The level $t = 0$ is the demand generation level or customer node level. Different demand locations produce different product types, which, in our case, are waste types. Each facility location, at depot level t, $t \neq 0$, can hold

all the given product types, in separate containers. The sizes of these containers constitute the set of size configurations available η^t, at that depot level t.

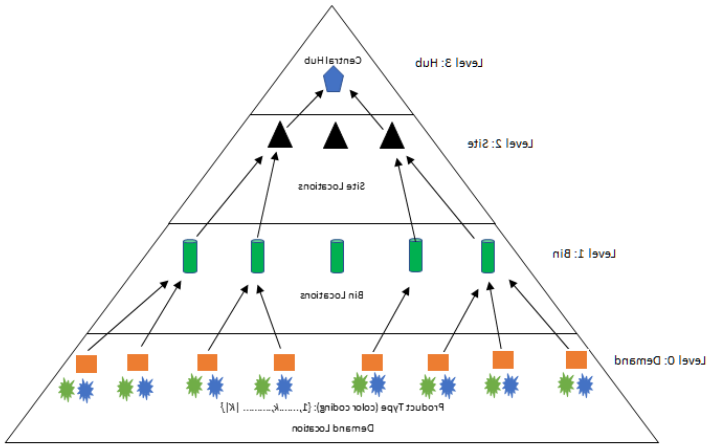


Figure 1: An illustration of hierarchy of a Waste Management System.

Next, we present a model formulation for the ME-LRP described above, designated as Model ME-LRP(MSW). The objective function of this model consists of two components. Based on the type of product being collected at an activated facility, the first component belongs to operating cost (p_{nk}^t), which is a function of the product type and the facility size, and constitutes the size configuration cost of the system. The second component belongs to the routing cost incurred by the vehicles traversing different nodes while collecting waste.

$$\text{Model ME-LRP(MSW): } min \sum_{\forall\, t\in\tau}\sum_{\forall\, i\in\nu^t}\sum_{\forall\, n\in\eta^t}\sum_{\forall\, k\in\kappa} p_{nk}^t * x_{ink}^t + \sum_{\forall\, t\in\tau}\sum_{\forall\, u\in\nu^t\cup\nu^{t-1}}\sum_{\forall\, v\in\nu^t\cup\nu^{t-1}}\sum_{\forall\, k\in\kappa} b_{uv}^t * h_{uvk}^t \quad (1)$$

Subject to

$$\sum_{\forall u\in\nu^{t-1}} y_{uik}^t \le \sum_{\forall n\in\eta^t} M_n^t * x_{ink}^t ; \forall i \in \nu^t, \forall\, k \in\ \kappa \quad (2_t)$$

$$\sum_{\forall\, u\in\nu^t\cup\nu^{t-1}} h_{uik}^t\ = \sum_{\forall\, n\in\eta^{t-1}} x_{ink}^{t-1}\ ; \forall\, i \in \nu^{t-1}, \forall\, k \in\ \kappa \quad (3_t)$$

$$\sum_{\forall\, u\in\nu^t\cup\nu^{t-1}} h_{iuk}^t\ = \sum_{\forall\, n\in\eta^{t-1}} x_{ink}^{t-1}\ ; \forall\, i \in \nu^{t-1}, \forall\, k \in\ \kappa \quad (4_t)$$

$$\sum_{\forall\, u\in\nu^{t-1}} h_{iuk}^t\ = \sum_{\forall\, u\in\nu^{t-1}} h_{uik}^t\ ; \forall\, i \in \nu^t, \forall\, k \in\ \kappa \quad (5_t)$$

$$h_{iuk}^t \le \sum_{\forall\, n\in\eta^t} x_{ink}^t\ ; \forall\, i \in \nu^t, \forall u \in \nu^{t-1}, \forall\, k \in\ \kappa \quad (6_t)$$

$$h_{uik}^t \le \sum_{\forall\, n\in\eta^t} x_{ink}^t\ ; \forall\, i \in \nu^t, \forall u \in \nu^{t-1}, \forall\, k \in\ \kappa \quad (7_t)$$

$$y_{uv}^t k \le V^t * h_{uvk}^t\ ; \forall\, u \in \nu^t \cup \nu^{t-1}, \forall v \in \nu^t \cup \nu^{t-1}, u \ne v, \forall\, k \in\ \kappa \quad (8_t)$$

$$\sum_{\forall u\in\nu^t\cup\nu^{t-1}} y_{iuk}^t - \sum_{\forall u\in\nu^t\cup\nu^{t-1}} y_{uik}^t = \sum_{\forall u\in\nu^{t-2}} y_{uik}^{t-1}\ ; \forall\, i \in \nu^{t-1}, \forall\, k \in\ \kappa \quad (9_t)$$

$$y_{iuk}^t = 0\ ; \forall\, i \in \nu^t\ , \forall\, u \in \nu^{t-1}, \forall\, k \in\ \kappa \quad (10_t)$$

$$y_{uvk}^t \ge 0\ ;\ u \in \nu^t \cup \nu^{t-1}\ , \forall v \in \nu^t \cup \nu^{t-1}, u \ne v, \forall\, k \in\ \kappa \quad (11_t)$$

$$z_{uvk}^t \le I^t * h_{uvk}^t\ ; \forall\, u \in \nu^t \cup \nu^{t-1}\ , \forall v \in \nu^t \cup \nu^{t-1}, u \ne v, \forall\, k \in\ \kappa \quad (12_t)$$

$$z_{iuk}^{t} = i * h_{iuk}^{t} \; ; \forall \; i \in \nu^{t} \; , \forall \; u \in \nu^{t-1}, \forall \; k \in \; \kappa \tag{13_t}$$

$$z_{iuk}^{t} = i * h_{iuk}^{t} \; ; \forall \; i \in \nu^{t} \; , \forall \; u \in \nu^{t-1}, \forall \; k \in \; \kappa \tag{14_t}$$

$$\sum_{\forall u \in \nu^{t} \cup \nu^{t-1}} z_{iuk}^{t} - \sum_{\forall u \in \nu^{t} \cup \nu^{t-1}} z_{uik}^{t} = 0 \; ; \forall \; i \in \nu^{t-1}, \forall \; k \in \; \kappa \tag{15_t}$$

$$z_{uvk}^{t} \geq 0 \; ; \; u \in \nu^{t} \cup \nu^{t-1} \; , \forall v \in \nu^{t} \cup \nu^{t-1}, u \neq v, \forall \; k \in \; \kappa \tag{16_t}$$

$$x_{ink}^{t} \in \{0,1\}; \; \forall \; i \in \nu^{t}, \forall n \in \eta^{t}, \forall \; k \in \; \kappa \tag{17_t}$$

$$h_{uv}^{t} \in \{0,1\} \; ; \; u \in \nu^{t} \cup \nu^{t-1} \; , \forall v \in \nu^{t} \cup \nu^{t-1}, u \neq v, \forall \; k \in \; \kappa \tag{18_t}$$

The Constraints (2_t)-(18_t) correspond to each depot level $t \in \tau$ of the ME-LRP system except for $t = 1$. For level $t = 1$, the decisions are the same as those for level $t > 1$. The only change is incurred in the expressions of Constraints 3_t, 4_t and 9_t. This is because for $t = 1$, $\nu^{t} = \iota$. The expressions for the rest of the constraints remain unchanged. Thus, rewriting Constraints (3_t), (4_t) and (9_t) for $t = 1$, we have:

$$\sum_{\forall \; u \in \; \nu^{1} \cup \iota} h_{ulk}^{1} \; = 1 \; ; \forall \; l \in \iota, \forall \; k \in \; \kappa \tag{3_1}$$

$$\sum_{\forall \; u \in \; \nu^{1} \cup \iota} h_{luk}^{1} \; = 1 \; ; \forall \; l \in \iota, \forall \; k \in \; \kappa \tag{4_1}$$

$$h_{uvk}^{1} + h_{vuk}^{1} \; = 1 \; ; \forall \; (u, v) \in \Delta, \forall \; k \in \; \kappa \tag{19_1}$$

$$\sum_{\forall u \in \nu^{1} \cup \iota} y_{luk}^{1} - \sum_{\forall u \in \nu^{1} \cup \iota} y_{ulk}^{1} = D_{(l,j)k} * h_{ljk}^{1} + D_{(j,l)k} * h_{jlk}^{1} \; ; \forall \; l \in \iota, \forall \; k \in \; \kappa \tag{9_1}$$

Constraints (2_t) represent the space constraints for each depot location and product type based on the size configuration selected at the location. Constraints (3_t) and (4_t) are, respectively, the In-degree and Out-degree constraints at each node i that belongs to lower depot level $t = 1$. Constraints (5_t) ensure flow of a vehicle leaving and returning at each activated depot. Constraints (6_t) and (7_t) ensure that a certain depot location i is on a Hamiltonian path only if that location is activated. Constraints (8_t)-(11_t) capture the subtour elimination constraints (SECs) based on the SECs for the TSP presented in [11]. Similarly, Constraints (12_t)-(16_t) form the fixed destination constraints (FDCs), which are adopted from [9] and [1]. Constraint (19_1) is an additional constraint for $t = 1$ that follows from the street wise collection for each street.

4. Column-generation-based approach

Our initial investigation revealed that the time taken for the solution of Model ME-LRP(MSW) by a direct application of a solver in each iteration is quite high, and it increases with iterations. Note that in view of Constraints (3_t)-(5_t), a feasible solution to the ME-LRP would partition the set of streets into non-overlapping sets that are dictated by the facility serving all the streets belonging to a specific partition. Similarly, facilities in every echelon apart from echelon T can be partitioned into non-overlapping sets, that are dictated by facility in the higher echelon. Thus, any feasible solution to the ME-LRP of the proposed framework, would exhibit single/multiple *crews*, each of which starts from echelon T, and includes the streets and facility nodes served by a facility from a higher echelon.

Based on this insight, we present a *set-partitioning*-based formulation for the ME-LRP. First, we derive the formulation for a single product, i.e., $k = 1$. A crew $c \in C$ is expressed as subsets of facility locations across all the echelons that cover a subset of the streets. Thus, $c \subseteq \; (\bigcup_{t \in \tau} \nu^{t} \;) \cup \iota$. A crew c is represented by a column vector of size $L + \sum_{t \in \tau} (I^{t} * N^{t})$. Considering a single product-based system, $k = 1$, for the

illustration described in Figure 1, the length of the column vector is 8 + (5*2 + 3*2 + 1*1) = 25, where we assume that $N^t = 1$ for the echelons corresponding to 'Bin' and 'Site' locations, and $N^t = 1$ for the 'Central Hub'. The elements of a column vector corresponding to a crew c are partitioned by echelons. For a crew $c \in C$, the element α_{it}^c of its corresponding column vector is equal to 1 if the location $i \in \nu^t$ is covered by the network expressed by crew c, and equal to 0 otherwise. Similarly, for the customer echelon $\alpha_{i0}^c = 1$, if the customer at location $i \in \iota$ is served by the crew, and 0 otherwise. For the echelons corresponding to the facility locations ($t > 0$), $\alpha_{it}^c = \sum_{n \in \eta^t} \alpha_{int}^c$. Thus, the column vector corresponding to a crew c entails the set of streets served by the crew, and the size and location of the facilities used to serve those streets. The cost of recruiting a crew includes cost of: (1) the size configuration of the facilities, and (2) routing the vehicles in establishing the crew's network.

A feasible solution to the ME-LRP would exhibit a single visit to each street for collection of waste and to the activated facility locations, by a vehicle that starts its trip from a facility in the next higher echelon. Since every vehicle departs and arrives to the same facility location, we can say that a street is served by a single facility in the higher echelon. A similar argument holds true for activated set of facilities for all $t \in \tau \setminus \{T\}$. Thus, a feasible solution to the ME-LRP can be constructed by choosing a set of crews $c \in C$, such that: (1) all the streets are served, (2) a customer is served by exactly one crew, and (3) the selected set of crews exhibit non-overlapping sets of facilities. Condition (3) is essential for correctly capturing the building (and size configuration) costs of the facilities. If a facility is used by two (or more) crews, the calculation of the system cost would not be straight forward as the cost of building that facility would be considered multiple times instead of once.

As noted above, a feasible solution to the ME-LRP can be constructed by choosing a set of non-overlapping crews that serve an independent set of streets, and the combined coverage by the selected crews fulfills the demand of each customer. Accordingly, we define a set-partitioning formulation for representing the ME-LRP labelled as $\Gamma_1(C)$, which is defined on a set of crews C using the binary decision variable Y_c, which equals 1 if crew $c \in C$ is selected. To define the master problem of the column generation (CG) methodology we relax the binary restriction on variable Y_c and use the following relaxed version of $\Gamma_1(C)$:

$$\big(R{:}\,\Gamma_1(C)\big){:}\, min \sum_{\forall\, c \in C} H^c \;*\; Y_c \tag{20}$$

$$\sum_{\forall\, c \in C} \alpha_i^0 * Y_c \;=\; 1; \forall\, i \in \iota \tag{21}$$

$$\sum_{\{\forall\, c \in C\}} \Big(\sum_{\{\forall\, n \in \eta^t\}} \alpha_{in}^t\Big) * Y_c \leq\; 1; \forall\, i \in \nu^t, \forall\, t \in \tau \tag{22}$$

$$0 \leq\; Y_c \leq\; 1;\; ; \forall\, c \in\; C; \forall\, i \in \nu^t, \forall\, t \in \tau \tag{23}$$

We relax the binary restrictions on variable Y_c as demonstrated in Constraints (23), which is due to the fact that when the number of columns is finite i.e., the size of C is finite, the solution to the LP relaxation of (Γ_1) would always be an integer solution due to the total unimodularity of the columns in C. Thus, solving the LP relaxation of (Γ_1) with a finite set C yields a feasible solution to (Γ_1), which is also an upper bound of the optimal solution to (Γ_1). For an optimal solution Y^* to (Γ_1), and dual variables γ_u and β_{ut} correspond to Constraints (21) and (22), respectively, the corresponding pricing problem can be expressed as follows:

$$\text{Pricing: ME-LRP}(Y^*){:}\, min \sum_{\forall\, t \in \tau} \sum_{\forall\, i \in \nu^t} \sum_{\forall\, n \in \eta^t} p_n^t * x_{in}^t \;+ \sum_{\forall\, t \in \tau} \sum_{\forall\, u \in \nu^t \cup \nu^{t-1}} \sum_{\forall\, v \in \nu^t \cup \nu^{t-1}} b_{uv}^t * h_{uv}^t \;-\; \sum_{\forall\, (u,v) \in \nu^1 \times \iota} \gamma_u * h_{uv}^1 \;- \sum_{\forall\, t \in \tau} \sum_{\forall\, (u,v) \in \nu^t \times \nu^{t-1}} \beta_{ut} * h_{uv}^t \tag{24}$$

subject to Constraints (2_t)-(18_t) $\forall\, t \in \tau$ and Constraints (19_1). The initial sets of columns are generated using the Nearest Neighbour Method for Capacitated Vehicle Routing Problem ([18]). This method is applied to generate routes for each echelon. The proposed heuristic for initial set of columns begins at $t = 1$ by finding routes, using Nearest Neighbour Method, that connects the nodes at $t = 0$, and assign them to facilities optimally by considering the assignment costs as the cost of connecting the facility to the initial and final nodes of the route. The selected facilities at $t = 1$ are treated as sources and the same heuristic is applied for finding routes connecting the sources and their assignments to facilities at $t = 2$. This process is carried out until $t = T$. This heuristic can also be used to generate an initial solution for Model ME-LRP(MSW) and Pricing Problem: ME-LRP(Y^*). The benefit of using this heuristic with the CG framework is that it enhances the propensity of the CG framework to produce a lower cost faster than using an optimization solver for Model ME-LRP(MSW).

5. Computational Investigation.

In this section, we present comparative results of solving various instances of the ME-LRP by direct application of a solver as well as the proposed CG method. The solver that we used is Gurobi version 10.0.2, implemented on an Intel Xeon E5-2620 v3 processor machine with 32 GB of RAM. The results are presented in Table 1. The method that reports the best performance for each instance with respect to (a) Upper Bound (UB), (b) Lower Bound (LB), and (c) Gap (in %) are marked in **bold**. We specified a computation time limit of 5,000 seconds. For both the methods, the initial solution/set of columns is generated using the heuristic proposed in the previous section. In the CG method, the pricing subproblem is solved as long as a negative objective function value is achieved provided it is at least 20% of the objective value of the corresponding $R\colon \Gamma_1(C)$ in magnitude. Along with this condition, if the pricing subproblem achieves a gap of 200% or less, the process is terminated, and new columns are added to set C. On the other hand, if the objective function value of the pricing subproblem is less than 20% of the objective value of the $R\colon \Gamma_1(C)$ in magnitude, we solve the pricing subproblem to full optimality.

The instances used for this study were the *.atsp instances from the TSPLIB collection [23]. In each instance, the nodes of the instance in consideration were divided into multiple echelons. For instances with number of nodes less than 150, the number of echelons, including the demand and facility level echelons, considered is 3, while 4 echelons were considered for instances with over 150 number of nodes. In case of 3 echelons, the nodes were approximately divided in the ratio of 7:5:2, while for the case of 4 echelons the nodes were divided in the ratio of 9:5:3:1. The suffix "_Low" and "_High", respectively, corresponds to the low and high ratio between the demand (which in this case, is the waste generated from a source) and the capacity of the transporting vehicle. The instances for which the value of UB is followed by an asterisk indicates that, after the initial heuristic solution, no new solution was found within 5,000 seconds. For all facility level echelons in each instance, the number of size configurations was 2, i.e., $N^t = 2, \forall t \in \tau$.

As indicated in Table 1, neither of the methods could solve the instances considered to optimality within 5,000 seconds (since UB ≠LB). In each instance, the CG method was observed to give a better UB value than that generated by Gurobi. As the size of instances increases, the difference between the UB of both the methods increases further, favoring the CG method. Instances with "_High" demand to vehicle capacity ratio were found to be more difficult to solve. Gurobi cannot solve instances ftv170, rbg323, and rbg358 while CG cannot solve rbg358 instance. The reason behind a better performance of CG can be attributed to reusing the heuristic in its pricing problem that enables it to achieve negative objective, thereby causing the CG method to find better solutions. Note that Gurobi obtains a better LB. However, the ability of the CG method to produce smaller UBs, especially for larger instances, enables it to obtain better results when implemented in a web-based tool, details of which are included in Appendix A2.

Table 1: Performance Comparison of the CG method vs the Gurobi Solver

Instance	Gurobi Solver[1]			CG Method[1]		
	UB	LB	Gap	UB	LB	Gap
ftv55_Low	22,148	**21,572**	**2.60%**	**22,033.5**	20,844.5	5.40%
ftv55_High	24,383	**22,780**	**6.57%**	**24,148**	21,032.6	12.90%
ftv64_Low	34,229.5	**32,996**	**3.60%**	**34,008**	32,765.5	3.65%
ftv64_High	41,222.5	**38,948.5**	**5.52%**	**41,110**	38,816.5	5.58%
ftv70_Low	25,758	**23,598.5**	**8.38%**	**25,579**	23,341	8.75%
ftv70_High	45,255	43,180	4.59%	**45,090**	**43,243**	**4.10%**
kro124p_Low	92,240.5	**76,426**	**17.14%**	**91,269**	74,812.6	18.03%
kro124p_High	113,265	**98,541**	13.00%	**110,953**	98,379	**11.33%**
ftv170_Low	70,922.5	**55,356.4**	21.95%	**65,509**	55,295.6	**15.59%**
ftv170_High	120,164.5*	**95,188.1**	20.79%	**117,236.5**	94,954.2	**19.01%**
rbg323_Low	97,431	**70,845.5**	27.29%	**95,826.0**	70,831.48	**26.08%**
rbg323_High	185,806.5*	154,257.0	16.98%	**184,016.49**	**154,311.72**	**16.14%**
rbg358_Low	114,109.5*	**79,317.0**	30.49%	**109,585.5**	79,296.3	**27.64%**
rbg358_High	**193,698***	153,815.5	**20.59%**	**193,698***	**153,816.2**	**20.59%**

6. Conclusion and Future Work

In this paper, we have presented a comprehensive framework for designing municipal solid waste (MSW) collection networks by integrating a Chinese Postman Problem–inspired formulation into a multi-echelon location-routing structure. The proposed ME-LRP(MSW) model captures the hierarchical nature of the MSW collection systems, where waste moves through a sequence of collection, transfer, and processing stages. By employing column-generation–based decomposition, the proposed approach significantly reduces computational burden and enables solution of realistic, large-scale urban instances and a smooth integration of this optimization engine into a web-based Solid Waste Echelon-Embedded Planner (SWEEP). Overall, this work establishes a scalable, analytically grounded, and practically deployable approach to MSW network design. There are several avenues for future research. First, the current framework considers recurring routing and operational costs but excludes capital investment decisions. Extending the formulation to incorporate joint location, sizing, and long-term amortization of facility infrastructure would broaden its applicability to strategic planning. Second, the model presented can be extended to include uncertainty arising from seasonal demand variation, traffic congestion, and operational disruptions. Stochastic or robust variants of the ME-LRP(MSW), as well as adaptive real-time optimization modules within SWEEP, would strengthen the reliability of the recommended designs.

[1] Both the Gurobi Solver and the CG method used the specified upper time limit of 5,000 seconds as none achieved optimality.

Appendix

A1: Definition of sets, parameters, variables and cost components for ME-LRP

Sets and Parameters

$K =$ number of different types of products.

$\kappa = \{1, \ldots, K\}$ is the set of different products available for pickup at targets (demand nodes) and an element of κ is represented by k.

$T =$ number of *depot level*s considered in the ME-LRP configuration. Note that the customer (target level is considered as level "0".

$\tau = \{1, \ldots, T\}$ is the set of *echelons*. Each $t \in \tau$ represents a facility level echelon. For example, $t = 1$ represents the level from where the vehicles are dispatched for the collection of materials from the target (customer) nodes; $t = 2$ represents the level from where the vehicles are dispatched for the collection of material from $t = 1$ level, and $t = T$ represents the terminal level.

$S =$ number of streets that need to be covered. Each street comprises of two terminal demand nodes.

$L =$ number of demand nodes. $2L = S$.

$\iota = \{1, \ldots., L\}$ is the set of target (demand) nodes where the products are generated/accumulated, and an element of ι is represented by l.

$\Delta = \{(u_1, v_1), (u_2, , v_2), \ldots, (u_S, v_S)\}$ represents the set of streets through their demand nodes where $\forall (u, v) \in \Delta \Rightarrow u, v \in \iota$.

$D_{(u,v)k} = D_{(v,u)k} \forall (u, v) \in \Delta$ represents volume of waste of type k generated from the street with (u, v) as its terminal points.

I^t number of potential facility locations in level t.

$v^t = \{1, \ldots, I^t\}$ is the set of facility locations in level t, and an element of v^t is represented by i.

N^t number of facility size configurations available for all product types in level t.

$\eta^t = \{1, \ldots, N^t\}$ is the set of the facility size configurations for all product types in level t, and an element of η^t is represented by n.

M_n^t is the capacity of configuration n for a facility in echelon t.

V_k^t is the vehicular capacity for transporting product type k, which is identical for all the vehicles that originate from level t, $t \neq 0$, for collecting the products from the nodes in $t - 1$ level.

Variables

Based on the type of decisions involved in the ME-LRP, we classify the variables into the following categories: location, size and routing.

Size Configuration Variables:

$x_{ink}^t \in \{0,1\}: = 1$ if size configuration n is selected for product type k at facility location i of echelon t, where $i \in v^t$, $t \neq 0$; $= 0$ otherwise.

Routing Variables:

$h_{uvk}^t \in \{0,1\}: = 1$ if an edge $u - v$, is used for routing a vehicle at echelon t, $t \neq 0$, collecting product type k from the nodes in echelon $t - 1$ $(u, v \in v^t \times v^{t-1}, u \neq v)$; $= 0$, otherwise.

$y_{uvk}^{t} \in$ R^{+}\$: Flow-based variable that captures flow on an edge $u - v$ to represent the routing of vehicles whose journey starts and terminates at a facility in echelon t, $t \neq 0$, in order to collect product type k from nodes in echelon $t - 1$ $(u, v \in v^{t} \times v^{t-1}, u \neq v)$.

$y_{uvk}^{t} \in$ R^{+}: Fixed Destination Constraints (FDC) flow variable for edge $u - v$ used for routing vehicles whose journey start and terminate at facilities in echelon t, $t \neq 0$, in order to collect product type k from nodes in echelon $t - 1$ $(u, v \in v^{t} \times v^{t-1}, u \neq v)$.

Cost Components:

b_{uv}^{t}: Cost of traversing edge $u - v$ by a vehicle that originates from depot level t, $t \neq 0$, and collects material from the nodes at level $t - 1$.

p_{nk}^{t}: Cost of operating a facility of size configuration n for product k in depot level t.

A2: Solid Waste Echelon-Embedded Planner (SWEEP)

SWEEP is a web-based decision support tool designed to optimize multi-echelon solid waste management networks. The system solves the Multi-Echelon Location Routing Problem (ME-LRP) by allowing users to define facility locations, vehicle capacities, and cost parameters across multiple echelons. It supports both interactive map-based input and bulk CSV uploads, generating execution plans that include detailed route tables, facility summaries, and interactive route visualizations. The tool provides detailed output in two forms: CSV files containing route tables, facility summaries, and cost breakdowns, and interactive route visualizations that display optimized vehicle routes on a map. For visualization, the application uses Leaflet for both input and output display, while ArcGis API (a product of ESRI) is employed for detailed route visualization with turn-by-turn directions and waypoint management as demonstrated in Figure 2.

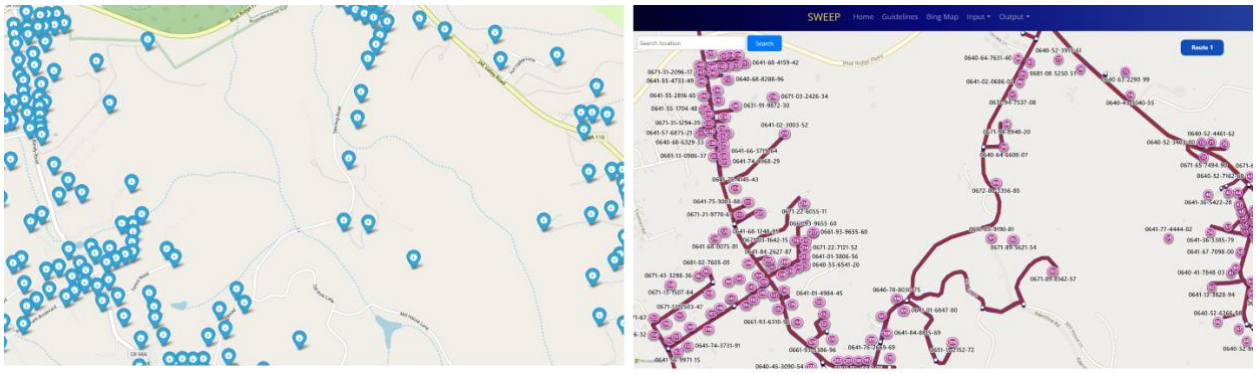


Figure 2: Visualization of an input (left) and the corresponding output routes (right) generated using SWEEP.

6.1 System Architecture

The system employs a modular three-tier architecture comprising a frontend interface, a backend API, and an optimization engine. These components are containerized using Docker to ensure consistent deployment and scalability.

Frontend Layer: Built as a Single-Page Application (SPA) using React 18 and Vite, the frontend manages user interactions and state via the React Context API. It utilizes Leaflet for facility placement and route visualization, with optional integration of the ArcGIS API for detailed turn-by-turn directions. The interface features dynamic filtering, allowing users to inspect optimization results by facility or echelon level.

Backend Layer: The backend is a RESTful service implemented in Flask (Python). It orchestrates data flow, serving endpoints for data submission, validation, and result retrieval. It handles file processing for CSV inputs and formats optimization outputs into JSON structures for the frontend and downloadable CSV reports for the user.

Optimization Engine: The core logic leverages the Gurobi Optimizer to solve ME-LRP using the proposed CG framework, where Gurobi is used to generate optimal solutions for (Γ_1) and pricing: ME-LRP(Y^*), the latter being accelerated using the Nearest Neighbour heuristic. It integrates with the ArcGIS API to calculate real-world distance matrices. The engine runs synchronously with a configurable time limit and optimality gap threshold, balancing solution quality with computational performance.

6.2 Data Flow and Integration

The system follows a unidirectional data flow pattern. User inputs—submitted via the interactive map or CSV upload—are aggregated into a structured JSON payload by the frontend. The backend validates this data and persists it to a temporary JSON file which serves as the input instance for the optimization engine.

Upon invocation, the Gurobi-based solver constructs the model, executes the optimization, and exports the results (facility selections, route assignments, and costs) to raw CSV files. A dedicated backend module transforms these raw outputs into consumed JSON formats and generates finalized route tables. The frontend then retrieves these results via GET endpoints to render visual markers and populate statistical summaries.

6.3 Deployment

To facilitate portability, the application is deployed via Docker Compose, utilizing a bridge network for internal communication. The frontend is served via an Nginx web server, while the backend runs in a Python container with necessary solver dependencies. This containerized approach supports deployment on varied infrastructure, including cloud-based Kubernetes clusters, ensuring separation of concerns between the user interface and the computational logic.

References

[1] Aguayo, M.M., Avilés, F.N., Sarin, S.C., Sherali, H.D., 2022. A Two-Index Formulation for the Fixed-Destination Multi-Depot Asymmetric Travelling Salesman Problem and Some Extensions. Informatica 33 (4), 671–692.

[2] Alshaikh, R., Abdelfatah, A., 2024. Optimization Techniques in Municipal Solid Waste Management: A Systematic Review. Sustainability 16 (15), 6585.

[3] Alumur, S., Kara, B.Y., 2007. A new model for the hazardous waste location-routing problem. Computers & Operations Research 34 (5), 1406–1423.

[4] Asefi, H., Lim, S., 2017. A novel multi-dimensional modeling approach to integrated municipal solid waste management. Journal of Cleaner Production 166, 1131–1143.

[5] Asefi, H., Lim, S., Maghrebi, M., others, 2015. A mathematical model for the municipal solid waste location-routing problem with intermediate transfer stations. Australasian Journal of Information Systems 19.

[6] Asefi, H., Lim, S., Maghrebi, M., Shahparvari, S., 2019. Mathematical modelling and heuristic approaches to the location-routing problem of a cost-effective integrated solid waste management. Annals of Operations Research 273 (1-2), 75–110.

[7] Asefi, H., Shahparvari, S., Chhetri, P., Lim, S., 2019. Variable fleet size and mix VRP with fleet heterogeneity in Integrated Solid Waste Management. Journal of Cleaner Production 230, 1376–1395.

[8] Badran, M.F., El-Haggar, S.M., 2006. Optimization of municipal solid waste management in Port Said - Egypt. Waste management (New York, N.Y.) 26 (5), 534–545.

[9] Bektaş, T., Gouveia, L., Santos, D., 2020. Compact formulations for multi-depot routing problems: Theoretical and computational comparisons. Computers & Operations Research 124, 105084.

[10] Filipiak, K.A., Abdel-Malek, L., Hsieh, H.-N., Meegoda, J.N., 2009. Optimization of municipal solid waste collection system: case study. Practice Periodical of Hazardous, Toxic, and Radioactive Waste Management 13 (3), 210–216.

[11] Gavish, B., Graves, S.C., 1978. The travelling salesman problem and related problems.

[12] Ghiani, G., Laganà, D., Manni, E., Triki, C., 2012. Capacitated location of collection sites in an urban waste management system. Waste management (New York, N.Y.) 32 (7), 1291–1296.

[13] Hannan, M.A., Hossain Lipu, M.S., Akhtar, M., Begum, R.A., Al Mamun, M.A., Hussain, A., Mia, M.S., Basri, H., 2020. Solid waste collection optimization objectives, constraints, modeling approaches, and their challenges toward achieving sustainable development goals. Journal of Cleaner Production 277, 123557.

[14] Heidari, R., Yazdanparast, R., Jabbarzadeh, A., 2019. Sustainable design of a municipal solid waste management system considering waste separators: A real-world application. Sustainable Cities and Society 47, 101457.

[15] Hirbod, F., Karimi, T., Mohammadnazari, Z., Rabbani, M., Aghsami, A., 2024. Municipal solid waste management using multiple disposal location-arc routing and waste segregation approach: a real-life case study in England. Journal of Industrial and Production Engineering 41 (1), 81–100.

[16] Hrabec, D., Kůdela, J., Šomplák, R., Nevrlý, V., Popela, P., 2020. Circular economy implementation in waste management network design problem: a case study. Cent Eur J Oper Res 28 (4), 1441–1458.

[17] Laporte, G., Nickel, S., Saldanha-da-Gama, F. Introduction to Location Science, in: , pp. 1–21.

[18] Liu, F., Lu, C., Gui, L., Zhang, Q., Tong, X., Yuan, M., 2023. Heuristics for Vehicle Routing Problem: A Survey and Recent Advances.

[19] Nawar, K.N., Mahbub, T., Tashfiq, R.A., Rashid, T.U. Municipal Solid Waste Collection, Transportation, and Segregation, in: , pp. 29–71.

[20] Ogwuleka, T.C., Agunwamba, J.C., 2009. Scheduling of solid waste collection routes. Nigerian Journal of Technology 28 (1), 79–92.

[21] Ortiz-Astorquiza, C., Contreras, I., Laporte, G., 2018. Multi-level facility location problems. European Journal of Operational Research 267 (3), 791–805.

[22] Rabbani, M., Heidari, R., Farrokhi-Asl, H., Rahimi, N., 2018. Using metaheuristic algorithms to solve a multi-objective industrial hazardous waste location-routing problem considering incompatible waste types. Journal of Cleaner Production 170, 227–241.

[23] Reinelt, G., 1991. TSPLIB—A Traveling Salesman Problem Library. ORSA Journal on Computing 3 (4), 376–384.

[24] Sadeghian Sharif, N., Pishvaee, M.S., Aliahmadi, A., Jabbarzadeh, A., 2018. A bi-level programming approach to joint network design and pricing problem in the municipal solid waste management system: A case study. Resources, Conservation and Recycling 131, 17–40.

[25] Sahoo, S., Kim, S., Kim, B.-I., Kraas, B., Popov, A., 2005. Routing Optimization for Waste Management. Interfaces 35 (1), 24–36.

[26] Samanlioglu, F., 2013. A multi-objective mathematical model for the industrial hazardous waste location-routing problem. European Journal of Operational Research 226 (2), 332–340.

[27] Sharholy, M., Ahmad, K., Mahmood, G., Trivedi, R.C., 2008. Municipal solid waste management in Indian cities - A review. Waste management (New York, N.Y.) 28 (2), 459–467.

[28] Terefe, A., 2015. Route optimization using arc routing for household waste collection.

[29] Wolfsberger, T., Pinkel, M., Polansek, S., Sarc, R., Hermann, R., Pomberger, R., 2016. Landfill mining: Development of a cost simulation model. Waste management & research : the journal of the International Solid Wastes and Public Cleansing Association, ISWA 34 (4), 356–367.

[30] Yousefloo, A., Babazadeh, R., 2020. Designing an integrated municipal solid waste management network: A case study. Journal of Cleaner Production 244, 118824.

[31] Zografros, K.G., Samara, S., 1989. Combined location-routing model for hazardous waste transportation and disposal. Transportation Research Record (1245).

Biography.

Akshat Kothyari is pursuing his PhD in the Grado Department of Industrial and Systems Engineering at Virginia Tech. His research interest are in production scheduling, logistics, and applied mathematical programming.

Shaunak Juvekar holds a Master's degree in Computer Science from Virginia Tech (2024) and is currently a Software Engineer at CNH Capital. His professional interests include application development, software architecture, and digital transformation.

Subhash C. Sarin currently holds the Paul T. Norton Endowed Professorship in the Grado Department of Industrial and Systems Engineering at Virginia Tech. He has made research contributions in production scheduling, sequencing, applied mathematical programming, and analysis of and designing algorithms for the operational control of manufacturing and logistics systems. He is recipient of several prestigious awards at the university, state, national, and international levels.

8th Conference on Production Systems and Logistics

An Experimental Analysis Of A Scaled Electromagnetic In-Space Launcher: A Multi-Stage Solenoid Accelerator For Nanosatellite Payload Deployment

Bhushan Lohar[1], Sai Raghava Pathuri[1], Sumedh Patil[2], John Wade[1], Mir Ali[1], Lance Outerbridge[1]

[1]*University of South Alabama, Department of Systems Engineering, Mobile, USA*
[2]*W.P. Davidson High School, Mobile, USA*

Abstract

Electromagnetic Space Launch systems (ESLs) offer transformative potential for delivering low-cost payloads, like CubeSats, or deploying space debris remediation systems. This paper details the development, iterative optimization, and testing of a scalable multi-stage ESL prototype powered by supercapacitors to support sustainable, solar-integrated operation. The primary objective was maximizing payload acceleration by systematically optimizing solenoid coil density and validating a multi-stage electronic control system. Systematic testing identified vital operational parameters. Propulsion required strict magnetic polarity alignment by configuring sled permanent magnets (North inward) and solenoid electromagnets (North inward) for attraction. Sufficiently ferrous solenoid cores were deemed essential for required magnetic flux. Furthermore, payload magnet length proved critical; overly long magnets saturated the Hall Effect sensors, preventing stage deactivation and inducing parasitic drag. During optimization, solenoid windings from 70 to 110 turns were evaluated using Vernier Video Analysis and Arduino timestamps. The primary scientific contribution of this work is the empirical identification of 80-turn winding density efficiency ceiling, yielding a maximum single-stage acceleration of 2.61 m/s². Applying this specification, a five-stage prototype was constructed. Final validation demonstrated consistent, successful operation, achieving a maximum payload acceleration of 5.47 m/s² and a 0.88 m/s top velocity. This research successfully validates an Arduino-controlled, 3D-printed ESL architecture, establishing key constraints and performance metrics to serve as a robust foundation for all future scaled systems.

Keywords

Electromagnetic Space Launch System (ESLs); CubeSats; Space Debris Collector; Electromagnetism; Payload Kinematics

1. Introduction

The concept of Electromagnetic Space Launch systems (ESLs) represents a profound paradigm shift in the fundamental methods used to access space. By harnessing the formidable physical forces underlying railgun and coilgun technologies, these systems have been meticulously engineered to achieve specific velocities tailored to a wide range of payload requirements. The coilgun, for example, demonstrates significant versatility by accelerating a 14-kg mass to 6 km/s and possesses the theoretical potential to be optimized to reach an astounding 11 km/s with strategic enhancements [1]. Simultaneously, the railgun system

DOI: https://doi.org/10.15488/21022
ISSN: 2701-6277

consistently achieves 6 km/s for payloads up to 14 kg, showcasing robust and reliable performance in controlled flight evaluations [2]. This research delves into the intricate design details of these ESL systems while highlighting their primary operational principles and performance capabilities. It further explores how these launchers can be reconfigured to accommodate larger payloads of up to 150 kg (SmallSats and CubeSats), enabling the precise deployment of scientific packages within circular orbits that are ideal for sustained aeroballistic studies and the support of space laboratories [3]. The ultimate effectiveness of these systems is not solely defined by their velocity achievements but also by the strategic integration of small rocket systems. These secondary stages facilitate necessary orbital insertion and overcome the persistent challenge of atmospheric drag, thereby ensuring the reliability of payload delivery into Low Earth Orbit (LEO) [4]. Furthermore, this discussion evaluates the significant comparative advantages of ESLs over conventional chemical rocket launchers (in-space), with a particular emphasis on potential cost efficiencies, increased launch availability, and the inherent scalability for emerging applications such as microsatellite and nanosatellite deployments. By enhancing access to the orbital environment through innovative launch technologies, ESLs stand poised to redefine the future of in-space exploration and commercial satellite operations, paving the way for unprecedented advancements in aerospace engineering and multilateral space situational awareness.

This electromagnetic launcher accelerates a payload sled along an aluminum extrusion rail using pairs of solenoids, which are constructed by coiling insulated, enameled copper wire around a ferrous core to maximize the electromagnetic field generated by an electric current. To prevent the payload's magnets from getting permanently trapped at the centre of the coil, a phenomenon known as the "pull-back" effect, a precise control circuit must cut the power at the exact moment of transit. As the sled approaches, a Hall Effect sensor detects the magnet's proximity and outputs a binary signal to an Arduino Uno, which then triggers a high-speed MOSFET switch to rapidly pulse a high-current burst from a supercapacitor bank. This precise timing maintains the payload's momentum, but a designer cannot simply add infinite coils to further increase speed; higher coil counts increase electrical resistance and inductance, causing the magnetic field to dissipate too slowly and creating parasitic drag that decreases overall acceleration.

Focusing on an iterative design-build-test cycle, this project tests methodologies to increase the operational efficiency of this controlled, chemical-propellant-free acceleration system. Similar electromagnetic technology is already in active use, such as on aircraft carriers to launch fighter jets [5]. Because it requires minimal energy overhead, this system is ideally suited for launching materials and resources into orbit from low-gravity and low-atmosphere environments like LEO, the lunar surface, or small planets. Furthermore, the technology presents significant space protection and humanitarian applications, such as acting as a high-velocity projectile, serving as a space debris collector [6], or launching autonomous Zipline drones to deliver critical medical supplies to remote regions during disaster responses [7]. Ultimately, these innovations offer a sustainable and versatile transportation pathway for both terrestrial and extra-terrestrial environments.

2. State of Art: The Evolution of Electromagnetic Space Propulsion

The development of ESLs shifts propulsion from volatile chemicals to pulsed magnetic fields, utilizing either railguns (sliding-contact Lorentz force) or coilguns (non-contact induction mass drivers) that offer a theoretically infinite lifespan for repeated, high-cycle space missions. The scientific foundation emerged in the 1970s when Gerard K. O'Neil proposed the propellant-free mass-driver to sustainably transport high-frequency 1-10 kg lunar payloads [8], culminating in MIT's 1977 Mass Driver I prototype that achieved 30g accelerations to prove the viability of sequential magnetic coil kinematic propulsion. By 1978, expanded research into superconducting "quenchguns" for high-velocity nuclear waste disposal further validated that coilguns could theoretically achieve orbital velocities [8]. Transitioning to Earth-to-orbit applications in the 1990s exposed massive power and scale challenges; a comprehensive analysis [9] concluded that

accelerating 100-1000 kg payloads to 1-6 km/s sub-orbital speeds requires 40-260 GJ of energy and tracks exceeding one kilometer. This highlighted the "acceleration-hardening" constraint, where shortening tracks demands extreme acceleration and heavily hardened satellite packages to survive G-forces, thereby shifting modern focus to multi-stage operations that distribute forces across hundreds of sequential coils to smooth the acceleration curve. Contemporary high-performance laboratory research includes the 3-coil Reconnection Electromagnetic Launcher (REL) that leverages complex magnetic flux reconnections to maximize stage kick up to 150m/s [10], alongside the lunar-specific Reluctance Linear Electromagnetic Launcher [11] that utilizes the moon's low gravity and lack of atmosphere to launch bulk materials in ways infeasible on Earth. While demonstrating higher performance limits, these systems rely on high-voltage industrial power grids and custom-machined components, creating significant barriers to entry for low-cost, scalable deployments.

3. System Concept and Methodology

3.1 Derivation of ESL Concept

The concept for this ESL prototype was derived from the need for sustainable, low $\boldsymbol{\Delta v}$ orbital maneuvering systems (such as CubeSat deployment or debris remediation) that eliminate reliance on volatile chemical propellants. In deriving the optimal architecture, two primary electromagnetic topologies were evaluated: railguns and coilguns. Railguns rely on sliding electrical contact between rails and an armature to generate Lorentz forces. However, prior knowledge dictates that this sliding contact inevitably causes plasma arcing and severe electrode erosion, rendering railguns unsuitable for long-term, unmaintained space deployment. Therefore, a multi-stage solenoid accelerator (coilgun**)** concept was selected. This architecture utilizes non-contact magnetic induction to propel a ferrous projectile, theoretically offering an indefinite operational lifespan. Furthermore, to address the known risk of high-G vibration failures in delicate nanosatellites during single-stage launches, a multi-stage staging concept was derived to flatten the acceleration curve and reduce instantaneous jerk.

3.2 Scientific Assumptions and Operational Environment

This system concept is engineered under specific assumptions regarding the LEO and vacuum environment:

- **Zero Drag and Heating:** The absence of atmospheric pressure means aerodynamic drag and convective heating are negligible. This scientific basis allows the design to prioritize pure electromagnetic force efficiency over aerodynamic shaping.
- **Solar-Regenerative Power:** The concept assumes continuous access to orbital solar arrays. This drove the selection of supercapacitors as the power medium, as they bridge the gap between the high-amperage burst requirements of the solenoids and the slow, steady recharge cycles of solar panels.

3.3 Governing Physics

The system is governed by the Lorentz force, where the magnetic field strength ($\boldsymbol{B}$) at the center of the solenoid is defined by $\boldsymbol{B = \mu_0 \frac{N.I}{L}}$. To prevent the parasitic "suck-back" effect, a known limitation where a coil remains energized and pulls the payload backward, the concept requires an active control loop to sever power at the geometric midpoint, ensuring exclusively forward-vector magnetic attraction.

3.4 Methodological Approach

The primary methodological objective was to systematically vary independent variables to identify the configuration yielding the highest acceleration. The experiment was structured in two distinct phases: Phase I (Winding Density) and Phase II (Payload Mass). While this section establishes other preliminary methodological approach, the exact setup and process for overarching concept is covered in later chapter.

3.4.1 Control Strategy and Timing Logic

Because manual control is impossible at high speeds, an automated active control loop was developed for proximity detection. The methodology utilizes a magnetic proximity sensor that outputs a LOW signal until detecting the payload's magnetic signature. This signal is inverted by an external microcontroller and a high-speed driver to trigger the gate of a low-side power MOSFET. Serving as a high-speed switch, the transistor applies current to pull the magnet forward and terminates it exactly as the payload reaches the center, allowing it to glide freely into the next stage.

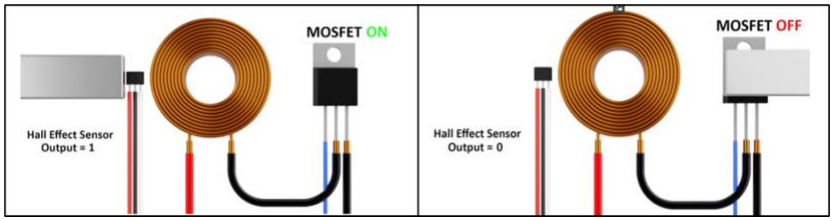


Figure 1: Solenoid Circuit Design Concept [12]

3.4.2 Thermal Management

A significant engineering challenge in this methodology is thermal management. Although the solenoid wire typically overheats under continuous current, the methodology dictates that it can safely handle massive amperage bursts since the sled clears the stage in a split second. Thermal loads are further managed by placing the MOSFET switches underneath aluminum heatsinks. To establish true physical constraints rather than timing inconsistencies, programmable circuit simulators were used to define strict pulse activation windows. As a redundant safety measure, the methodology includes a programmed time limit that places the launcher into standby mode if the sled becomes lodged.

4. Experimental Setup

4.1 Structural Framework & Sled Design

The structural framework utilized an aluminum extrusion secured with T-nuts to serve as the primary guide rail. Custom 3D-printed brackets were fabricated to mount the solenoids and sensors, while the control circuitry (finalized with 1k–10k ohm pull-down resistors and a momentary safety switch) and MOSFETs were positioned on the opposite side. The 3D-printed payload sled featured a superior fiducial tracking marker and lateral notches for neodymium magnets. Empirical findings dictate these magnets must face North-pole inward for correct electromagnetic interfacing [14].

4.2 Hardware Specifications & Solenoid Construction

Electrical propulsion was driven by a specific 16.2V 83F supercapacitor bank. Standardized solenoid cores comprised one-inch steel bolts, nylon washers, and steel nuts inside 0.75-inch PVC sleeves. These were clearance-verified, secured in the custom 3D brackets with high-strength adhesive, and hand-wound with 0.8 mm (16-gauge) enameled copper wire while maintaining five-inch power leads.

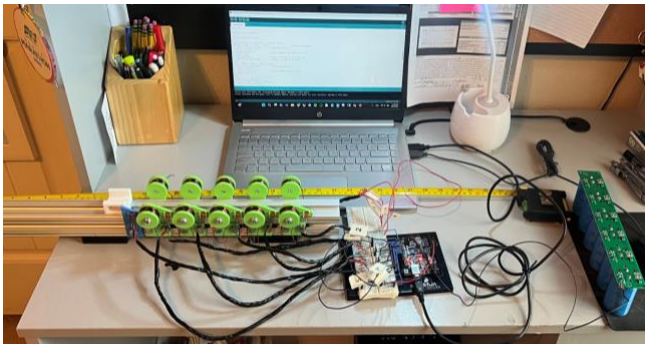


Figure 2: Trial Setup of ELS [14]

The physical control circuitry integrated A3144 Hall Effect sensors for proximity detection, routed to an Arduino control board. The logic was enhanced by TC4428A high-speed power MOSFET drivers triggering BOJACK IRF740 N-Channel (10A 400V) and IPP026N10NF2SAKM MOSFETs to handle the extreme current surges (which exceeded 100 amps during launch bursts, well above the 0.8 mm wire's 5-amp continuous rating). Initial functionality testing utilized a one-second pulse via the momentary switch to verify magnetic polarity; any backward sled movement indicated that the coil's electrical leads needed reversing.

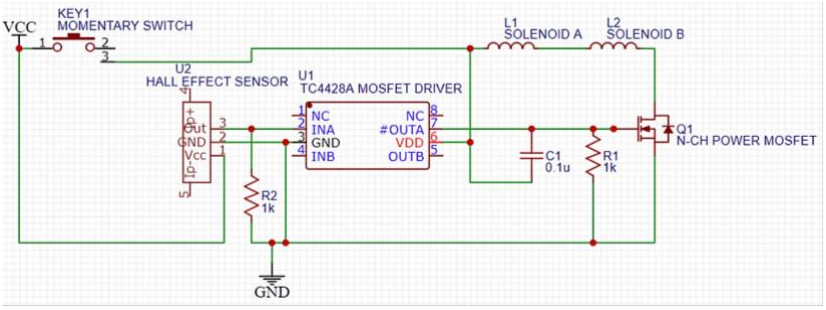


Figure 3: Power Discharge Circuit Design Concept [8]

The primary engineering objective of this research was to systematically vary payload mass and solenoid winding density to identify the configuration yielding the highest acceleration. Employing an agile development process, the experiment was divided into two distinct phases to isolate variables and maximize cumulative efficiency:

- **Phase I: Winding Density:** The team focused on identifying the optimal number of solenoid windings through a series of six distinct trials. The number of turns on each solenoid bilateral pair was systematically varied to include configurations of 110, 100, 90, 80, and 70 turns [13]. Three trials per pair were conducted to guarantee balanced propulsion and statistical significance.
- **Phase II: Payload Mass Variation:** Once the three most efficient coil counts were identified, the team expanded the launcher to a full 10-solenoid configuration. In this phase, the independent variable was the payload mass. Using standard lab test weights, the team increased the sled's mass in increments of 10 grams (from 20g to 60g) to observe the effect of inertia on electromagnetic propulsion efficiency [14].

For final high-precision trials, a Digilent programmable circuit simulator and LabVIEW software were used to generate exact 200ms to 300ms pulses, establishing the optimal activation window. Throughout testing, a recording device coupled with a laptop running Vernier Video Analysis and Vernier Graphical Analysis software tracked the fiducial marker at 29.97 frames per second. This allowed for the generation of detailed position-time graphs, determining peak velocity and acceleration via the velocity curve's maximum instantaneous slope. Operators mitigated risks from high-current discharge and hypervelocity transients using ANSI Z87.1 ocular protection and Class 00 dielectric gloves.

5. Results and Discussion

To evaluate coil efficiency, the team utilized a multi-stage digital workflow, importing high-resolution trial recordings into video analysis. By tracking a high-contrast fiducial marker frame-by-frame and isolating X-axis displacement, the team generated precise motion data. These datasets were exported as GAMBL files to Vernier Graphical Analysis to determine granular kinematic characteristics. The team applied a linear regression model ($y = mx + b$) to the velocity-time graphs; the resulting slope ("m") provided the constant acceleration (m/s^2) for each trial. These values were cross-referenced with internal Arduino timestamps to ensure MOSFET switching aligned with peak physical acceleration. By comparing results across the 110 to 70-turn trials, the team identified the specific point where electrical resistance began to impede magnetic flux. This empirical analysis successfully identified the most efficient launcher configuration, fulfilling the project's primary engineering objective.

Table 1: Velocity Trials of Solenoids with 110 to 70 Turn

	Velocity			
Number of Coils	Trial 1 (m/s)	Trial 2 (m/s)	Trial 3 (m/s)	Average (m/s)
110	0.259	0.286	0.26	0.268
100	0.247	0.281	0.248	0.259
90	0.152	0.16	0.156	0.156
80	0.334	0.332	0.343	0.336
70	0.235	0.243	0.237	0.238

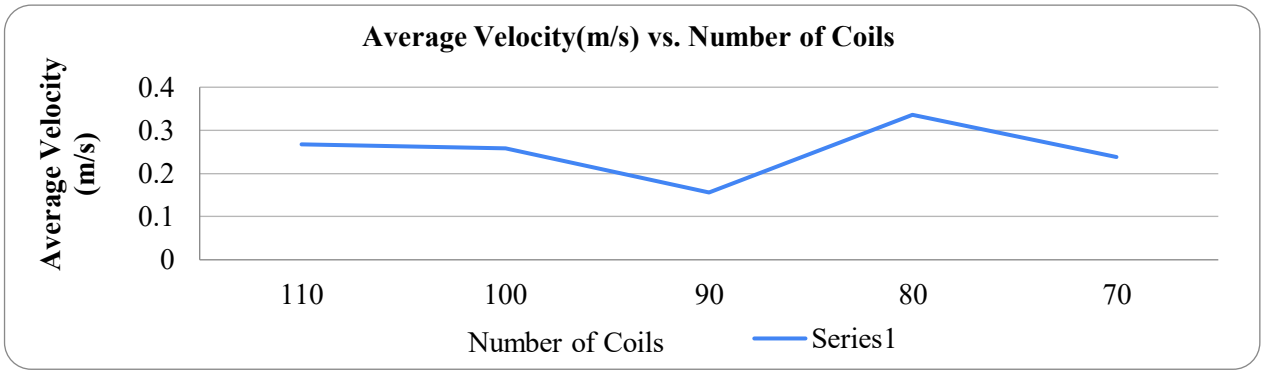


Figure 4: Average Velocity(m/s) vs. Number of Coils

Table 2: Acceleration Trials of Solenoids with 110 to 70 Turn

	Acceleration			
Number of Coils	Trial 1 (m/s^2)	Trial 2 (m/s^2)	Trial 3 (m/s^2)	Average (m/s^2)
110	2.26	2.07	2.17	2.17
100	2.12	2.43	2.14	2.23
90	1.1	1.15	1.18	1.14
80	2.47	2.66	2.69	2.61
70	1.84	2.04	2.03	1.97

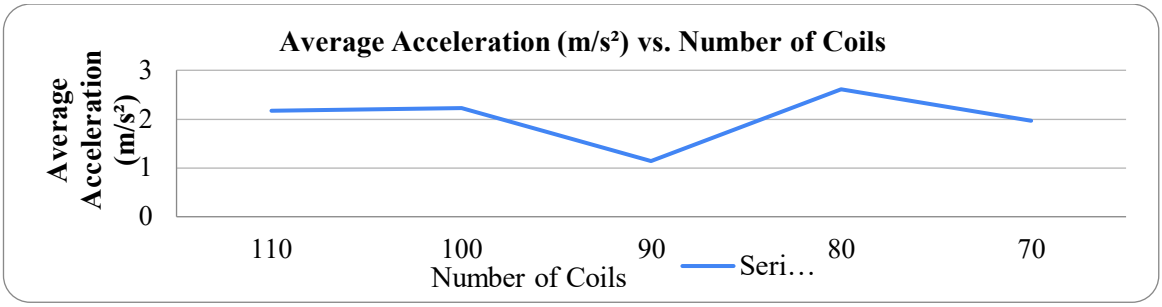


Figure 5: Average Acceleration (m/s²) vs. Number of Coils

Upon identifying 80 turns as the optimal winding density for maximum acceleration, the propulsion system was scaled by fabricating four additional solenoid pairs. Initial empirical testing of a single 80-turn coil pair yielded a peak acceleration of 2.61 m/s². Despite the expansion to a five-stage configuration, achieving the theoretical target of 11 m/s² remains a significant engineering challenge given current hardware constraints. Based on the performance curves of the existing prototype, it is projected that the current multi-stage architecture can sustain a maximum realized acceleration of approximately 5 m/s².

Table 3: 80 Coils Velocity and Acceleration T test Comparison with other Coils

T Condition	T test (Velocity)	T test (Acceleration)
80 Coils vs 70 Coils	0.0006%	0.0003%
80 Coils vs 90 Coils	0.1968%	0.0013%
80 Coils vs 100 Coils	0.1678%	0.5303%
80 Coils vs 110 Coils	0.0049%	0.6456%

To ensure kinematic precision, the team utilized a dual-method analytical framework. The primary methodology involved high-speed photogrammetry via Vernier Video Analysis, where frame-by-frame tracking generated position-time and velocity-time profiles. These datasets were then processed in Vernier Graphical Analysis to derive average and peak acceleration values. A secondary, independent data stream was captured via Arduino Serial Monitor timestamps. By correlating discrete Hall Effect sensor trigger intervals with physical sensor displacement, the team applied fundamental kinematic derivatives.

$$v = \frac{\Delta d}{\Delta t}; a = \frac{\Delta v}{\Delta t}$$

This cross-referencing revealed that consistent acceleration occurred in specific stages, emphasizing the necessity of multi-tool analysis for capturing complex electromagnetic transients.

Our empirical results validate the theoretical constraints identified in [9] regarding the acceleration limits. The novel scientific contribution of this paper is that while high-voltage systems like the REL [10] achieve velocities up to 150m/s, our prototype demonstrates that for low-power, solar-compatible logistics, the efficiency ceiling is dictated by resistance-flux trade-off at 80 turns. The results also confirm that a small-scale COTS based electromagnetic launcher is a feasible solution for the rapid acceleration of payloads via sequential magnetic stages. This technology holds transformative potential for extraterrestrial operations, particularly in low-gravity and low-atmosphere environments such as the lunar surface for precision maneuvers. By utilizing ESL systems to launch materials or equipment into orbit, space agencies can achieve significant fuel savings compared to traditional chemical rockets. Furthermore, the system's ability to reach a precise, desired velocity makes it an ideal candidate for the orbital insertion of CubeSats from manned satellites or space stations.

Beyond payload deployment, the team identified several innovative applications for high-velocity electromagnetic systems:

- **Space Debris Remediation:** A collector system, such as a specialized net, could be attached to the sled and launched at high velocities to intercept orbital debris. The kinetic energy provided by the launcher would allow the net to cover vast orbital distances before its trajectory decays, facilitating safer near-Earth environments.
- **Deep Space Exploration:** The principle of high-velocity initial launch can be applied to interplanetary probes and rovers, providing the necessary momentum for deep space transit without the excessive mass of a chemical first-stage booster.
- **Interplanetary Logistics:** Reusable and energy-efficient, these systems offer greater acceleration for small-scale parcels, significantly reducing the overhead associated with frequent logistical missions between orbital platforms.

Despite these clear advantages, future iterations must address critical challenges, including significant large-scale power storage requirements, the necessity of post-launch secondary propulsion for precise orbital maneuvering, and the need for high-strength, lightweight payload materials to withstand destructive rapid acceleration forces. Testing observations dictate that future research must prioritize reducing vehicle friction as mechanical drag on the aluminum rail occasionally hindered sliding and improving coil efficiency. Specifically, the current 200-millisecond energizing pulse is inefficiently long; firing five sequential coil sets over a half-meter track in one second limits hypervelocity potential, meaning this non-linear acceleration sequence must be drastically shortened. Consequently, future studies will utilize advanced simulations and precise timing adjustments to determine whether equidistant or progressively spaced solenoids best maximize the acceleration profile.

6. Conclusions and Future Work

This research successfully validated a multi-stage solenoid accelerator prototype for ESLs, but more importantly, it established a novel empirical baseline for low-powered ESL systems. By identifying 80 coils of 21-gauge enameled copper wire per solenoid as the optimal winding density, this study provides a crucial engineering constraint for future researchers designing sustainable, supercapacitor-based orbital manoeuvring systems. A finalized five-stage assembly utilizing these 80-coil pairs achieved a 0.88 m/s maximum velocity and 5.47 m/s² peak acceleration, confirming the efficacy of sequential electromagnetic staging. Operationally, the system effectively converted magnetic force to kinetic energy via a low-friction 3D-printed sled with neodymium magnets gliding along an aluminum extrusion rail. However, testing revealed two operational challenges: the initial magnetic pulse occasionally failed to overcome the sled's stationary inertia at the first electromagnet, and thermal buildup risks prevented consistent use of Hall Effect sensors. To mitigate the thermal risks, a computer-controlled interface was used to manage power distribution in place of standard MOSFET logic.

To address current limitations, future research will maximize electromagnetic acceleration via strictly orderly, unidirectional solenoid coiling and inter-winding insulating layers to concentrate the magnetic field. To eliminate rail contact friction, the team will integrate a maglev-inspired opposing-pole suspension system. Structurally, current wooden brackets will be replaced with unified 3D-modeled housings that precisely align the Hall Effect sensor, MOSFET, and aluminum heatsink directly beneath the solenoid, mitigating data irregularities caused by manual assembly. Once the sled consistently clears each electromagnet's midpoint, the system will revert to a fully autonomous MOSFET-driven circuit. Refining these mechanical and electrical constraints will advance the prototype toward a scalable, reliable solution for low-gravity payload deployment and space debris remediation.

References

[1] Schroeder, J.M., Gully, J.H., Driga, M.D., 1989. Electromagnetic launchers for space applications. IEEE Transactions on Magnetics. 25 (1), 504–507. https://doi.org/10.1109/20.22590.

[2] McNab, I.R., 2018. Electromagnetic space launch considerations. IEEE Transactions on Plasma Science. 46, 3628–3633.

[3] Kou, B., Li, L., Zhang, C., 2009. Analysis and optimization of thrust characteristics of tubular linear electromagnetic launcher for space-use. IEEE Transactions on Magnetics. 45, 250–255.

[4] Miller, L.A., Rice, E.E., Earhart, R.W., Conlon, R.J., 1984. Preliminary analysis of space mission applications for electromagnetic launchers.

[5] Li, W., Yang, T., Xin, Y., 2024. Concept of an Auxiliary System for Carrier-Based Aircraft Catapult. IEEE Transactions on Applied Superconductivity, 34(3), 1-5.

[6] Micheli, D., Santoni, F., Giusti, A., Delfini, A., Pastore, R., Vricella, A., Albano, M., Arena, L., Piergentili, F., Marchetti, M., 2017. Electromagnetic absorption properties of spacecraft and space debris. Acta Astronaut. 133, 128–135. https://doi.org/10.1016/j.actaastro.2017.01.015.

[7] Demuyakor, J., 2020. Ghana go digital agenda: The impact of Zipline drone technology on digital emergency health delivery in Ghana. Shanlax Int. J. Arts Sci. Humanit. 8, 242–253. https://doi.org/10.34293/sijash.v8i1.3301.

[8] Snow, W. R., Kolm, H. H., 1992. Electromagnetic launch of lunar material. NASA, Johnson Space Center, Space Resources, Vol. 2: Energy, Power, and Transport.

[9] Turman, B.N., Lipinski, R.J., 1995. Electromagnetic coilgun launcher for space applications. Albuquerque, New Mexico. University of North Texas Libraries, UNT Digital Library.

[10] Fan, G., Wang, Y., Xu, Q., Nie, X., Zhongming, Y., 2019. Design and analysis of a novel three-coil reconnection electromagnetic launcher. IEEE Trans. Plasma Sci. 47 (1), 814–820. https://doi.org/10.1109/TPS.2018.2874287.

[11] Einat, M., Orbach, Y., 2023. A multi-stage 130 m/s reluctance linear electromagnetic launcher. Sci. Rep. 13 (1), 218. https://doi.org/10.1038/s41598-022-27022-z.

[12] Kwan, E., n.d. Electromagnetic Launcher Prototype. Instructables. Available at: https://www.instructables.com/Electromagnetic-Launcher-Prototype/ (accessed 16 December 2025).

[13] Patil, S., Lohar, B., Thomas, T., Outerbridge, L., 2024. Attempting increasing the efficiency of an electromagnetic launcher. https://doi.org/10.13140/RG.2.2.34392.61440.

[14] Patil, S., Lohar, B., Ali, M., Thomas, T., Outerbridge, L., 2025. Electromagnetic space launcher – Alabama Science and Engineering Fair (ASEF) 2025. https://doi.org/10.13140/RG.2.2.14882.75209.

Biography

Dr. Bhushan Lohar is an Assistant Professor in Systems Engineering and founder of Advanced Systems Research Laboratory (ASRL) at the University of South Alabama (USA). He is a faculty advisor of the INCOSE student chapter and USA Satellites Club at the USA and President of the INCOSE Blues Chapter. He has over 10 years of experience combining the automobile and aerospace industry, working on systems engineering, research & development, project management, and quality management. His research interests include Model-Based Systems Engineering, Model-Based Patterns, Space Systems Engineering, Artificial Intelligence and Digital Twins.

Sai Raghava Pathuri is a Ph.D. student in Systems Engineering at the University of South Alabama, where he is conducting research at ASRL under Dr. Bhushan Lohar. His current research focuses on the development and modeling of complex systems such as Uncrewed Systems (UxS) integrated with Digital Twins. His research has significant implications for disaster response applications.

Sumedh Patil is a senior at W.P. Davidson High School in Mobile, Alabama, currently completing the rigorous IB Diploma program. As the lead designer for the school's championship robotics team, he possesses a strong foundation in engineering and technical design. Sumedh intends to major in aerospace engineering, with the long-term goal of contributing to the space industry. His research interests are centered on electromagnetism, CubeSats, and the establishment of extraterrestrial colonies.

Dr. John Wade is a fire protection engineer and partner at Gulf States Engineering, Inc. Dr. Wade is also an adjunct professor in Systems Engineering at the College of Engineering at the University of South Alabama. He has 30 years of experience in the design of mechanical and fire protection systems for heavy industry. His research interests include fire protection systems, project engineering, decision analysis, and firefighting tactics. Dr. Wade holds a patent for advanced wildfire-fighting systems.

Mir Ali is Model-Based Systems Engineer with 11+ years across aerospace, IT, and manufacturing. He uses CAMEO SysML to connect requirements, behavior, and hardware for complex, safety-critical systems. His focus is clean system architecture, clear requirements, and practical test and validation plans aligned with NASA, ISO, and MIL standards. He likes taking messy problems, structuring them, and driving them through integration, verification, and sustainment with cross-functional teams.

Lance Outerbridge is a Management Systems Specialist at the University of South Alabama's College of Engineering, where he supports complex operations and technical systems. He is currently pursuing a Master of Science in Systems Engineering at the university. Lance has extensive experience in process control and safety engineering, with expertise in configuring DCS and PLC platforms such as Rockwell, Honeywell, and DeltaV, as well as developing SCADA and HMI systems to optimize integrated operations.

8th Conference on Production Systems and Logistics

Implementing Lean Six Sigma Methodologies in an Industry 4.0 Manufacturing Setup: A Case Study

Shantall Cisneros Saldana, Heike Markus
Hof Univeristy of Applied Sciences, Hof, Germany

Abstract

The integration of Lean Six Sigma (LSS) principles with Industry 4.0 technologies offers an effective way to improve operational performance in modern manufacturing environments. This case study presents the systematic application of LSS within an Industry 4.0 manufacturing laboratory, employing the DMAIC (Define, Measure, Analyze, Improve, Control) framework to identify and eliminate inefficiencies. In a context dominated by automation, cyber-physical systems, and data-driven decision making, the study addresses persistent challenges including inaccurate sensor readings, limited user knowledge of advanced systems, and suboptimal material storage configurations. Through a combination of statistical analysis, root-cause identification, and iterative process redesign, the intervention resulted in measurable improvements in process reliability, user experience, and overall workflow efficiency. The findings highlight how LSS can serve as an effective bridge between traditional continuous-improvement methodologies and digitalized manufacturing operations. Moreover, the outcomes position the integration of LSS as a foundational enabler for Industry 5.0, where human-centricity, resilience, and sustainability will increasingly shape manufacturing system design.

Keywords

Lean Six Sigma; Industry 4.0; Industry 5.0; DMAIC; Cyber-Physical Systems; Smart Manufacturing

1. Introduction

The new era of Industry 4.0, characterized by advanced automation, interconnected systems, and data-driven decision-making in manufacturing, represents a transformative manufacturing landscape[1,2]. At the core of this transformation lies the systematic use of industrial data, where real-time, high-volume, and high-velocity data streams enable predictive, adaptive, and self-optimizing production systems[3]. These systems, devices, machines, production sites, and products can communicate and exchange information with each other, establishing a smart factory. The core of the Industry 4.0 structure is the Cyber-Physical System (CPS) [4], which provides high levels of monitoring, transparency, and control at all levels of production systems [5]. CPS are physical objects equipped with integrated systems, sensors, radio frequency identification (RFID) readers, and actuators. These sensors generate high-resolution, real-time data that enable unprecedented insights into process performance settings. Such data availability creates the foundation for what is referred to as Lean Six Sigma 4.0 (LSS 4.0), where classical continuous-improvement methodologies are enhanced through industrial IoT, big data analytics, and digital visualization tools[6]

The transition from traditional manufacturing to Industry 4.0 highlights new socio-technical complexities: increased system automation does not automatically eliminate errors, user confusion, or technical bottlenecks. This study employs data-driven Lean Six Sigma to investigate these issues, illustrating how

DOI: https://doi.org/10.15488/21023
ISSN: 2701-6277

structured problem-solving improves operational performance, user experience, and environmental responsibility within a smart factory lab. In Lean Six Sigma 4.0, data is not supportive but becomes a strategic asset that drives hypothesis testing, root-cause identification, and statistically sound decision making [6]

In the early 1990s, lean production principles gained widespread recognition [7]. Six Sigma methodology seeks to continuously reduce process variation and eliminate defects, improving process performance and quality [8]. The merger of Lean and Six Sigma resulted in Lean Six Sigma (LSS) [9], which focuses on eliminating waste and variation through the DMAIC framework.

The LSS methodology aims to improve processes, enhance customer satisfaction, and improve financial performance through systematic elimination of defects and waste [8]. In a lean manufacturing facility, improving flow and reducing non-value-added activities leads to reduced waste and lower costs [10]. The integration of Industry 4.0 and LSS enables smarter, more efficient manufacturing processes [11]. Recent studies emphasize that LSS 4.0 reduces dependence on manual data collection by leveraging machine connectivity and automated data acquisition, thereby strengthening methodological rigor and scalability [6,10].

DMAIC is the standard Six Sigma problem-solving method, with phases for Define, Measure, Analyze, Improve, and Control [12]. Digital manufacturing settings can use DMAIC to analyze cycle times, equipment states, and failure modes by using sensor data, log files, and event traces. Timestamps from sensors allow precise calculation of processing times, forming the basis for statistical evaluation. Despite this potential, empirical work explicitly linking DMAIC phases to specific Industry 4.0 technologies such as RFID-based tracking, PLC logs, and digitally supported SOPs, remains relatively sparse. There is a need for structured case studies that show how these data streams are integrated, how measurement systems are validated, and how statistical tools validate the effectiveness of proposed improvements.

Industry 5.0 extends the primarily technology-driven focus of Industry 4.0 towards human-centric, sustainable, and resilient socio-technical systems [1]. It highlights the important role of humans in collaboration with smart machines, and the creation of systems that promote well-being, learning, and flexibility. In this context, LSS can act as a key enabler by framing improvements not only in terms of throughput and defects, but also in terms of cognitive load, ergonomics, safety, and environmental impacts such as waste and energy consumption [13].

This research implements Lean Six Sigma methodology in an Industry 4.0 lab setup with three primary research questions:

RQ1: How does the integration of data-driven Lean Six Sigma within an Industry 4.0 cyber-physical system (CPS) operationalize technical stability, quantified through Cp/Cpk, DPMO, and sigma levels, to serve as a foundational enabler for Industry 5.0 pillars?

RQ2: How does the implementation of data-driven Lean Six Sigma 4.0 facilitate social sustainability by enhancing user experience, reducing cognitive load, and institutionalizing knowledge management through intuitive collaboration between humans and machines?

RQ3: How can data driven Lean Six Sigma improve environmental sustainability via waste reduction, energy efficiency, cycle-time optimization?

Following this introduction, the article is structured according to the DMAIC framework, detailing the baseline measurement of the cyber-physical system, the implementation of data-driven improvements, and a final validation of the operational and sustainability outcomes.

2. Case Study: Industry 4.0 Manufacturing Setup

This case study explores an Industry 4.0 manufacturing setup at Hof University of Applied Sciences. The setup produces two types of chocolate boxes with lids: red and blue. The vertical integrated system comprises four distinct operational layers: customer web store, Enterprise Resource Planning (ERP) system, Manufacturing Execution System (MES), and the cyber-physical system (CPS) on the shop floor level.

The CPS comprises a control station, pallet storage station, handling station, and press station, as illustrated in Figure 1. Each station is controlled by a programmable logic controller (PLC). The shop floor uses 54 sensors to continuously monitor production progress and identify problems. These include color sensors, proximity sensors, induction sensors, RFID readers, and actuators.

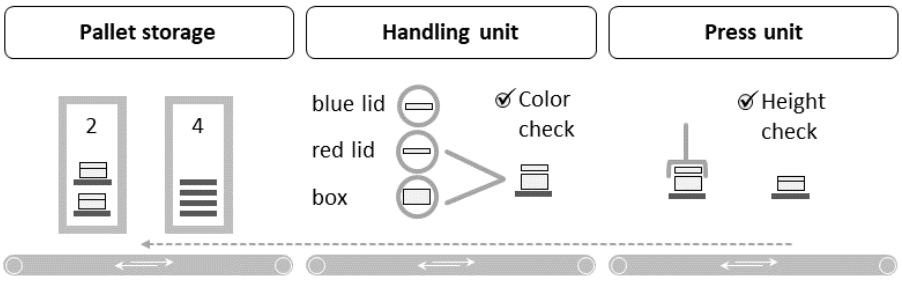


Figure 1: Cyber-Physical System at Hof University laboratory

This study adopts the DMAIC method in the CPS environment. During the defined phase, two approaches were used: visual inspection to identify areas for improvement, and sensor data collection to establish quantifiable baseline metrics. A total of 55 production orders were performed to collect baseline data. Through visual assessment of the cyber-physical system performance, some key issues were identified, including the lack of complete operating instructions. The CPS currently lacks clear and detailed step-by-step descriptions to guide the activities required to achieve optimal system operational performance. This deficiency poses a significant challenge, especially for new users.

2.1 Baseline Analysis: Measure and Define Phase

A total of 55 production orders (POs) were processed within the CPS, allowing for a detailed examination of cycle-time performance and process stability. The analysis identified the handling station as the critical bottleneck in the production flow. For this workstation, the upper specification limit (USL) for cycle time was set at 75 seconds, consistent with system design parameters and operational performance requirements.

The dataset showed a process with significant variation and distinct performance levels. For all production orders, the handling station's average cycle time was 67.1 seconds, with a 120.1-second standard deviation, leading to a 179.1% variation coefficient. Cycle times varied between 31 and 934 seconds, with a median of 46.0 seconds. Under ideal conditions, the system operated with notable stability: approximately 76.4% of all orders were completed within 44–48 seconds, with a standard deviation of only 1.2 seconds. Based on these findings, the handling station is dependable and consistent in its operation.

A second group of cases (14.5% of POs), lasting between 53 and 76 seconds, reflected intended test scenarios designed to mirror common operator or system induced disturbances. Minor delays, usually 20–30 seconds long, were caused by various factors including lid problems, incorrect color choices, brief material shortages, and misplaced pallets. These events show genuine deviations that occur often. Other disruptions were observed in 3.6% of production orders, with cycle times between 104 and 131 seconds. These delays, resulting from a shortage of materials like lids and boxes, required operators to stop and manually restock.

In the most extreme situation, which was 1.8% of the sample, a production order took 934 seconds due to a critical system failure. This incident occurred due to RFID tag misalignment, which halted processing and

required full manual intervention. Moreover, a small fraction of outliers (3.6% of orders) showed cycle times of roughly 31 seconds, indicating incomplete process cycles.

2.2 Process Capability Metrics (Baseline)

Process capability indices quantify how well a process meets specification limits. The Cpk index [14] accounts for process centering and spread. Using the handling station upper specification limit of 75 seconds and n=55:

$$C_{pk} = \frac{USL-\mu}{3\sigma} = \frac{75-67.1}{3\times120.1} = 0.022 \quad (1)$$

Where:

USL= 75 seconds (handling station target); $\bar{x}$ (mean) = 67.1 seconds; σ (std dev) = 120.1 second

The Cpk formula expresses how many standard deviations (3σ) fit between the process mean (μ) and the upper specification limit (USL); a very low Cpk (0.022) indicates the process is poorly centered and highly variable, making it incapable of consistently meeting the specification. The baseline capability assessment is shown in Table 1. The standard deviation sample (s) was calculated using the following formula:

$$s = \sqrt{[\Sigma(xi - \bar{x})^2 / (n-1)]} \quad (2)$$

Where:

s= sample standard deviation; $\sum(x_i - \bar{x})^2$=sum of squared deviations of each observation from the mean; n-1= degrees of freedom

Table 1: Baseline capability assessment

Metric	Value	Interpretation
Mean handling time	67.1 s	Process centered near USL; high risk of exceeding limit
Standard deviation	120.1 s	High variability; process not in statistical control
Cpk	0.022	(target ≥ 1.33); negative centering risk
Non-conforming (>75 s)	4/55 (7.3%)	One-in-fourteen orders exceed specification
DPMO	72,727	Corresponds to ~1.9σ performance
Sigma level	1.9σ	below the Six Sigma target (6σ ≈ 3.4 DPMO)

The completion of 76% of orders in 44-48 seconds validates the core process's stability and capability. The remaining variability arises not from equipment limitations but from predictable disturbances in human–system interaction and system design. Operator knowledge gaps, such as failing to replenish materials, created avoidable delays, which were amplified by the absence of clear standard operating procedures. Lid misalignment, improper material choice, and pallet errors happened when users used their own methods, with no directions. Limited visual feedback further hindered operators' ability to detect misalignments or material shortages in real time. Design elements were also a factor: big storage vessels enabled parts to move, causing issues with jams and falls. The Lean Six Sigma analysis therefore reframed these issues not as operator errors but as solvable design and procedural weaknesses. This perspective aligns with Industry 5.0's human-centric approach, which emphasizes supporting users through clearer procedures and more intuitive system design.

3. Analysis and Improvement

Using Five-Why analysis, a representative disturbance was examined to reveal its design flaw. The lid was incorrectly seated in the robotic gripper, which caused the jam. During the relocation, the lids inside the storage vessel were displaced. This shift was enabled by the vessel's oversized geometry; its dimensions exceeded the necessary tolerances, allowing excessive lateral movement. The oversizing resulted from conservative safety factors used in the initial design, where exact precision wasn't outlined. This showed a deeper organizational problem: not enough teamwork between mechanical design and process control engineering. Ultimately, the root cause was the vessel geometry's mismatch with the box and lid dimensions, which created conditions for systematic mishandling. As a result, the improvement needed a design-focused strategy, which called for carefully designed storage containers to limit material movement and allow dependable robotic grasping.

Sensor-data analysis uncovered a strong correlation between order duration and the number of items processed per orders can be seen in Figure 2. Orders containing a single item required an average of 101 seconds, whereas orders with two or more items averaged 68 seconds, a reduction of approximately 33 seconds, or nearly 40%. This pattern suggests that the handling and press stations operate more efficiently when setup and verification times are amortized over multiple items. The dominance of verification time in single-item orders guided subsequent improvement efforts toward minimizing operator-dependent confirmation steps and enhancing storage and visual-management design.

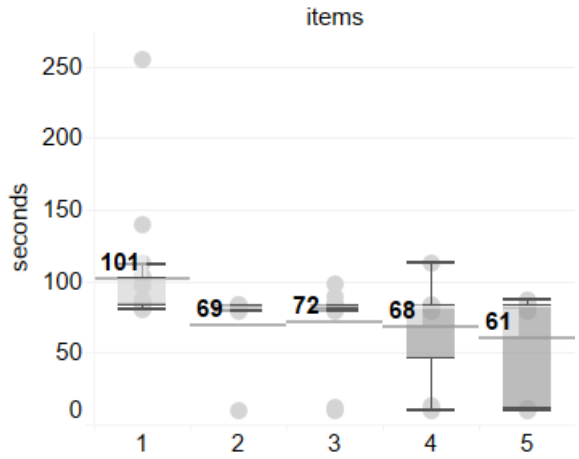


Figure 2: Handling Station Cycle Time by Number of Items per Order

3.1 Improvement Solutions

3.1.1 Standard Operating Procedures (SOP) Development

A major source of variation stemmed from the absence of clear, standardized operational guidance, particularly for new users. Operators often devised their own strategies for material checks and pallet alignment, resulting in inconsistent execution and avoidable errors. To address this, a visual SOP was developed, covering pallet verification, material availability checks, gripper alignment, and sensor confirmation prior to operation. The SOP is displayed at the control station, and new operators are required to verbally confirm each step. This intervention aimed to reduce knowledge-dependent variation, ensure consistent material preparation, and prevent improvisation-driven deviations.

3.1.2 Visual Poka-Yoke for RFID Alignment

The 934-second delay traced back to incorrect pallet placement, which caused RFID misalignment and system rejection. To prevent recurrence, high-visibility markers were applied to indicate the correct pallet orientation relative to the RFID tag. This simple visual poka-yoke assisted the new operators in achieving

correct placement before initiating automated processing, eliminating a critical failure mode and reducing reliance on manual intervention.

3.1.3 Precision Storage Vessel Redesign

The original storage vessels, designed with generous safety margins, were 5–10 mm larger than the required dimensions. This misfit allowed lids and boxes to shift during robotic handling, resulting in falls and jams. Newly designed and 3D-printed containers with tolerances of ±0.5 mm now match the exact geometry of the materials, as illustrated in Figure 3. The redesign included anti-tipping lips. This modification helped to eliminate misalignment events, enhance robotic precision, and reduce cycle-time variability.

Figure 3: Storage vessel before and after improvement

4. Control Phase.: Validation and Post Improvement Analysis

To assess the effectiveness of the implemented improvements, a validation experiment was conducted using 20 newly executed production orders and 2 new users. Every improved system setting was in use during this trial. The optimized CPS logic, which eliminated redundant RFID reads, was fully deployed; the redesigned 3D-printed precision storage vessels replaced the original oversized containers; operators followed the newly established SOP checklist; visual poka-yoke markers indicating correct pallet orientation were applied; and the visual material-replenishment protocol ensured proactive stocking of boxes and lids. This setup allowed for a fair assessment of the changes in real-world scenarios.

Performance at the handling station improved after the intervention. The mean cycle time decreased from 67.1 seconds in the baseline condition to 47.3 seconds, representing a 19.8-second (29.5%) reduction. The range of values became much smaller: the standard deviation went from 120.1 to 8.2 seconds, and the coefficient of variation went from 179.1% to 17.3%. Cycle times were strictly kept between 45 and 56 seconds; no orders took longer than 75 seconds. Additionally, there were no unusual occurrences, such as lack of materials, lid issues, jams, or RFID problems. The handling process operated with complete conformance during the validation period, demonstrating the stabilization achieved through design, procedural, and control-logic improvements.

Using the unchanged upper specification limit (USL) of 75 seconds and n=20, the post-improvement capability index increased substantially. The process's Cpk was roughly 1.13, calculated as:

$$C_{pk} = \frac{USL-\mu}{3\sigma} = \frac{75-47.3}{3\times 8.2} = 1.13 \quad (3)$$

This result shows progress from a deficient process (Cpk = 0.022) to one nearing statistical control, with performance around 2.0–2.5 sigma, as can be seen in Table 2. Even though it's below the benchmark, the improvement is significant, especially with process variation nearly gone. DPMO dropped from 72,727 to about 6,000, which is a 91.7% decrease, and all previous special causes were gone.

Table 2: Baseline capability post-improvement

Metric	Baseline	Post-Improvement	Change	Status
Mean	67.1 s	47.3 s	−19.8 s (−29.5%)	Improved
Std Dev	120.1 s	8.2 s	−111.9 s (−93.2%)	Major reduction
CV	179.1%	17.3%	−161.8 pp (−90.4%)	Stabilized
Cpk	0.022	1.13	+1.108 (5000% improvement)	2.0σ
DPMO	72,727	~6,000	−66,727 (−91.7%)	Near 2σ target
Non-conforming %	7.3%	0%	−7.3 pp (−100%)	Eliminated test cases
Special causes	10 events	0 events	−100%	Process stabilized

A comparison of the baseline and post-improvement cycle-time distributions can be seen in Figure 4. The baseline data display wide dispersion, with extreme values ranging from 30 to 934 seconds and numerous clusters of elevated cycle times. In contrast, the post-improvement data form a narrow band centred around 47 seconds, with no observations below 45 seconds or above 56 seconds. The disappearance of both delays, such as the 934-second RFID failure and moderate test-case disturbances, highlights the robustness of the redesigned system.

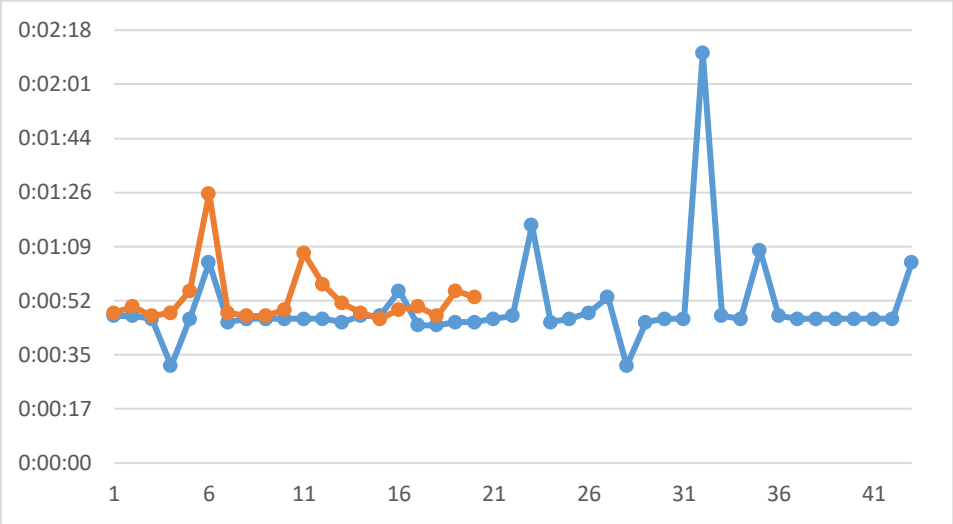


Figure 4: Handling Station Cycle Time Distribution – Before (Blue) and After (Red)

A two-sample t-test was performed to determine whether the observed improvement in mean handling time was significant. Using the baseline (n=55) and post-improvement (n=20) datasets, the test yielded a t-statistic of approximately 2.15 with a corresponding p-value below 0.01. These results show that the reduction in cycle time is highly unlikely to be due to chance. The post-improvement process shows not only a significantly lower mean but also a substantial decrease in variability, validating the effectiveness of the implemented solutions through both practical and statistical evidence. A limitation of this study is the sample size disparity between the baseline (n=55) and the validation phase (n=20). The limited scope of this post-improvement sample, with only two new operators, suggests the findings are a 'proof-of-concept'. Further research with larger, more diverse groups is needed to validate scalability in high-volume industrial setting.

5. Discussion

The case study shows a shift from attributing variability to operator error toward recognizing underlying design and system-interaction challenges. Events such as lid falls, incorrect color selection, material shortages, and RFID misalignment were consistently linked to ambiguous layout, insufficient visual

guidance, or imprecise material design rather than just individual mistakes. By reframing disturbances as solvable design issues, the improvement strategy aligns with Industry 5.0's human-centric vision, in which systems support operators through intuitive interfaces, error-proofing, and reduced cognitive load. The introduction of SOPs, visual indicators, precision storage, and poka-yoke mechanisms not only stabilized the process but also enabled new users to achieve flawless execution, strengthening competence and operational confidence.

The integration of CPS data streams with LSS methodology significantly enhanced analytical depth by enabling detection of millisecond-scale inefficiencies and rare disturbances, such as the 934-second RFID failure, that would not have been captured through manual observation. This transparency improved measurement adequacy supported capability analysis, and accelerated identification of special causes, while LSS tools such as capability indices, DPMO, hypothesis testing, and structured root-cause analysis converted extensive CPS-generated data into actionable insights. The DMAIC framework guided the conversion of insights into enduring design and procedural improvements.

The project's improvements support Industry 5.0 goals of energy and material efficiency. Reducing cycle times by 29.5% improves energy efficiency impact. Quantitative thermal analysis of this specific lab setup reveals a system-wide heat loss of 80.3 W, which translates to an energy loss of 0.32 kWh over four hours of operation [13]. At a rate of 0.0032 kWh per box, the optimization of cycle times and the elimination of 10 special-cause disturbances directly prevent unnecessary energy-intensive idle time and mechanical friction. Since the redesigned material handling system eliminated material drop-offs, the post-improvement operation also achieved better material utilization.

6. Conclusion

The integration of Lean Six Sigma, cyber-physical systems, and Industry 5.0 principles delivered substantial and measurable improvements. This study's primary contribution is showing how LSS 4.0 acts as a governance framework, converting automated data into sustainable design solutions, thus connecting technical optimization with human-centered needs. This case study directly addresses all research questions: RQ1 demonstrates operational excellence through Cpk improvement from 0.022 to 1.13, DPMO reduction (72,727→6,000), and 29.5% cycle-time gains, operationalizing Industry 5.0's human-centric, resilient, and sustainable pillars. RQ2 confirms social sustainability via enhanced user experience (new operators succeed first try), zero safety incidents (lid falls eliminated), and knowledge management through standardized SOPs. Data transparency reduced cognitive load and supported learning-by-doing, reinforcing the role of data-driven LSS 4.0 in human-centric system design. Regarding RQ3, the application of the DMAIC framework contributed to environmental sustainability by eliminating special-cause disturbances and further shortening cycle times by 29.5%.

This case provides empirical evidence that LSS and LSS 4.0 can operationalize Industry 5.0 principles. It shows how data-driven analysis reveals design-solvable challenges, how CPS transparency enhances statistical precision, and how human-centric design supports both performance and operator well-being. This approach proves that even in a laboratory environment, focusing on human-centric design and reduced cognitive load is a foundational requirement for modern smart factories.

Future work should incorporate long-term Statistical Process Control (SPC) to monitor sustainment. Extending the DMAIC methodology to other operational layers, such as ERP and MES systems, would enable a more holistic assessment of performance and structured integration of operator feedback to support continuous improvement. Scaling the framework to other CPS-based learning factories would enable broader validation of its generalizability.

In conclusion, integrating Lean Six Sigma into the Industry 4.0 lab achieved operational excellence and sustainability while paving the way toward Industry 5.0. It strengthens human–technology collaboration,

promotes user well-being, and sustainable practices. This case study provides insight into the future-ready mindset needed to thrive in modern manufacturing that embraces both Industry 4.0 and the emerging horizons of Industry 5.0. Future-oriented systems must foster synergistic growth between people and technology while developing the future skills required to lead ongoing industrial transformation.

References

[1] Ghobakhloo, M., Mahdiraji, H.A., Iranmanesh, M., Jafari-Sadeghi, V., 2024. From Industry 4.0 Digital Manufacturing to Industry 5.0 Digital Society: a Roadmap Toward Human-Centric, Sustainable, and Resilient Production. Inf Syst Front.

[2] Weyer, S., Schmitt, M., Ohmer, M., Gorecky, D., 2015. Towards Industry 4.0 - Standardization as the crucial challenge for highly modular, multi-vendor production systems. IFAC-PapersOnLine 48 (3), 579–584.

[3] Pongboonchai-Empl, T., Antony, J., Garza-Reyes, J.A., Komkowski, T., Tortorella, G.L., 2024. Integration of Industry 4.0 technologies into Lean Six Sigma DMAIC: A systematic review. Production Planning & Control 35 (12), 1403–1428.

[4] Davies, R., Coole, T., Smith, A., 2017. Review of Socio-technical Considerations to Ensure Successful Implementation of Industry 4.0. Procedia Manufacturing 11, 1288–1295.

[5] Hofmann, E., Rüsch, M., 2017. Industry 4.0 and the current status as well as future prospects on logistics. Computers in Industry 89, 23–34.

[6] Arcidiacono, G., Pieroni, A., 2018. The revolution lean six sigma 4.0. International Journal on Advanced Science, Engineering and Information Technology 8 (1), 141–149.

[7] Kolberg, D., Zühlke, D., 2015. Lean Automation enabled by Industry 4.0 Technologies. IFAC-PapersOnLine 48 (3), 1870–1875.

[8] Salah, S., Rahim, A., Carretero, J.A., 2010. The integration of Six Sigma and lean management. International Journal of Lean Six Sigma 1 (3), 249–274.

[9] Cherrafi, A., Elfezazi, S., Govindan, K., Garza-Reyes, J.A., Benhida, K., Mokhlis, A., 2017. A framework for the integration of Green and Lean Six Sigma for superior sustainability performance. International Journal of Production Research 55 (15), 4481–4515.

[10] Sanders, A., Elangeswaran, C., Wulfsberg, J., 2016. Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing. JIEM 9 (3), 811.

[11] Skalli, D., Charkaoui, A., Cherrafi, A., Garza-Reyes, J.A., Antony, J., Shokri, A., 2023. Industry 4.0 and Lean Six Sigma integration in manufacturing: A literature review, an integrated framework and proposed research perspectives. Quality Management Journal 30 (1), 16–40.

[12] Sokovic, M., Pavletic, D., Pipan, K.K., 2010. Quality improvement methodologies–PDCA cycle, RADAR matrix, DMAIC and DFSS. Journal of achievements in materials and manufacturing engineering 43 (1), 476–483.

[13] Saldana, S.C., Kapoor, A., Acharya, S., Markus, H., 2025. Energy Efficiency Analysis in Industry 4.0 set up leveraging Industry 5.0 methods for Sustainable Manufacturing: Case Study. Procedia Computer Science 253, 594–602.

[14] Boyles, R.A., 1991. The Taguchi capability index. Journal of quality technology 23 (1), 17–26.

Biography

Shantall Cisneros Saldana is a Research Associate and PhD candidate at Hof University of Applied Sciences. She holds a degree in Petroleum and Natural Gas Engineering and a Master's in Operational Excellence. Her professional and academic work focuses on Industry 4.0, Industry 5.0 and supply chain analytics, contributing to sustainable, human-centered digital transformation initiatives. Her doctoral research focuses on the systematic development and validation of Industry 5.0 Future Skills aligned with industrial competence needs.

Prof. Dr. Heike Markus is Professor at Hof University of Applied Sciences and Head of the Supply Chain Digitization research group. She has extensive experience in strategic planning and supply chain management at Munich Airport and in global industrial firms. She currently leads the Master's programs in Operational Excellence and Digital Transformation and conducts research in Industry 4.0 and 5.0, applied analytics, Data Science and Artificial Intelligence.

8th Conference on Production Systems and Logistics

Optimizing Last-Mile Distribution in an Emerging-Economy Bottled Water Firm Through VRP and DDMRP Integration

Boris Castañeda-Rios, Veralucía Zamalloa-Peralta, Marcos Fernando Ruiz-Ruiz and Juan Carlos Quiroz-Flores

Carrera de Ingeniería Industrial, Universidad de Lima, Lima, Perú.

Abstract

This study examines the operational limitations of a bottled water distribution company, a recurring issue identified in previous research on distribution planning and route optimization. The company faced urgent challenges related to a low On-Time-In-Full (OTIF) performance of 62%, high logistics costs, and a significant proportion of incomplete deliveries. To address these limitations, the research proposed an integrated model based on the Vehicle Routing Problem (VRP) and Demand Driven Material Requirements Planning (DDMRP), aimed at reorganizing order allocation and improving synchronization between production and distribution activities. The model was evaluated through a comparative analysis of As-Is and To-Be scenarios using discrete-event simulation. Results showed an increase in the OTIF indicator to 69%, a reduction in incomplete deliveries, and a 20% decrease in late orders. Additionally, the approach generated estimated annual savings of PEN 142,443, achieved a positive Net Present Value (NPV) of PEN 238,981, and increased labor productivity by 33%.

Keywords

Last-Mile Distribution; Vehicle Routing Problem (VRP); Demand-Driven MRP (DDMRP); Operational Performance; Emerging Economies.

1. Introduction

Small and Medium Enterprises (SMEs) represent more than 99% of formal companies in Peru and play a fundamental role in national economic development [1]. Within this context, the bottled water market has experienced sustained growth, with consumption exceeding 689 million liters in 2023 [2]. Despite this expansion, several institutional and academic reports indicate the persistence of operational deficiencies among SMEs, including a high dependence on manual planning processes, low levels of digitalization, and limited coordination between operational areas [3].

These structural limitations directly affect the performance of distribution activities. Previous studies report that manual route planning leads to excessive travel distances and recurrent delivery delays, negatively impacting punctuality and operational reliability [4]. Additional research highlights that non-optimized routing decisions increase operational times and reduce fleet utilization efficiency [5]. Moreover, inconsistencies in order assignment have been shown to generate operational overloads and disorganized delivery sequences, further deteriorating distribution performance [6]. The absence of adequate geographic segmentation has also been identified as a factor that restricts the efficiency of last-mile distribution operations [7].

DOI: https://doi.org/10.15488/21024
ISSN: 2701-6277

Internal coordination issues between production, storage, and dispatch processes further intensify these challenges, generating shortages and frequent order reprogramming that compromise product availability [8]. In this regard, planning methodologies such as Demand Driven Material Requirements Planning (DDMRP) have demonstrated positive results in stabilizing inventory levels through the use of dynamic buffers [9]. Empirical evidence from Peruvian SMEs also reports reductions in stockouts following the implementation of this approach [10]. Based on this context, the present article proposes an integrated model combining the Vehicle Routing Problem (VRP) and DDMRP to improve distribution planning and internal synchronization in a bottled water company located in Metropolitan Lima.

2. State of the art

2.1 Last-Mile Distribution

Last-mile distribution is increasingly recognized as a critical stage in logistics, where customer preferences and delivery flexibility influence route efficiency and service performance [11]. In bottled water and similar fast-moving consumer goods, demand variability and high delivery frequency intensify operational disruptions and expose supply chain vulnerabilities [12]. Nevertheless, existing studies primarily emphasize routing performance without explicitly addressing synchronization between distribution planning and upstream inventory availability in SME contexts.

2.2 Vehicle Routing Problem (VRP)

The Vehicle Routing Problem (VRP) is a central framework for optimizing distribution networks by integrating capacity, distance, and service constraints [13]. Recent developments explore multi-echelon configurations to improve cost efficiency and service levels [14]. However, most VRP applications operate independently from inventory management systems, limiting their capacity to mitigate service failures caused by production–distribution misalignment.

2.3 Demand-Driven MRP (DDMRP)

DDMRP extends traditional MRP systems by incorporating decoupling points and dynamic buffers to improve material flow under uncertainty [15]. Research highlights its capacity to stabilize inventory levels and reduce variability compared to push-based systems [16], [17]. However, limited evidence examines its interaction with distribution-routing decisions in urban SME environments.

2.4 Inventory Optimization

Modelling and simulation approaches have proven effective for assessing the impact of DDMRP on service indicators and stock behavior [18], while integrating MRP-type systems with programming tools enhances responsiveness in SMEs [19]. Despite these advances, routing optimization and inventory stabilization remain predominantly studied as independent decision domains. VRP research focuses on delivery efficiency [13], [14], whereas DDMRP literature concentrates on buffer positioning and variability reduction [15]–[17]. As a result, limited empirical evidence exists on integrated frameworks that simultaneously synchronize route planning and inventory availability in SME distribution systems operating in emerging economies.

This study addresses this gap by proposing and validating an integrated VRP–DDMRP model supported by discrete-event simulation, applied to a bottled water SME in Metropolitan Lima and evaluated through operational and financial indicators.

2.5 Methodology

This study adopts an applied case-study design aimed at diagnosing operational inefficiencies and validating an integrated improvement model within a real SME distribution context. The research process was structured into three phases: diagnosis, model development, and validation. In the diagnostic stage,

operational data—including OTIF performance, delivery times, incomplete orders, and inventory records—were analyzed to identify structural deficiencies in routing practices and production–distribution coordination.

A structured literature review, summarized in Appendix 1, supported the selection of decision-support tools according to contextual feasibility criteria such as route efficiency, demand synchronization, material availability, and implementation constraints typical of SMEs. Based on this evaluation, complementary methodologies were selected to address both external distribution inefficiencies and internal inventory instability. The proposed model was subsequently validated through discrete-event simulation, comparing baseline (As-Is) and improved (To-Be) scenarios under controlled experimental conditions.

2.6 Proposed Model

The proposed model was developed to address the operational deficiencies identified in the 20-liter bottled water distribution process, particularly low OTIF performance, delivery variability, and misalignment between production planning and dispatch execution. As illustrated in Figure 1, the model is structured into three sequential components: diagnosis, implementation, and validation, which together form an integrated improvement framework.

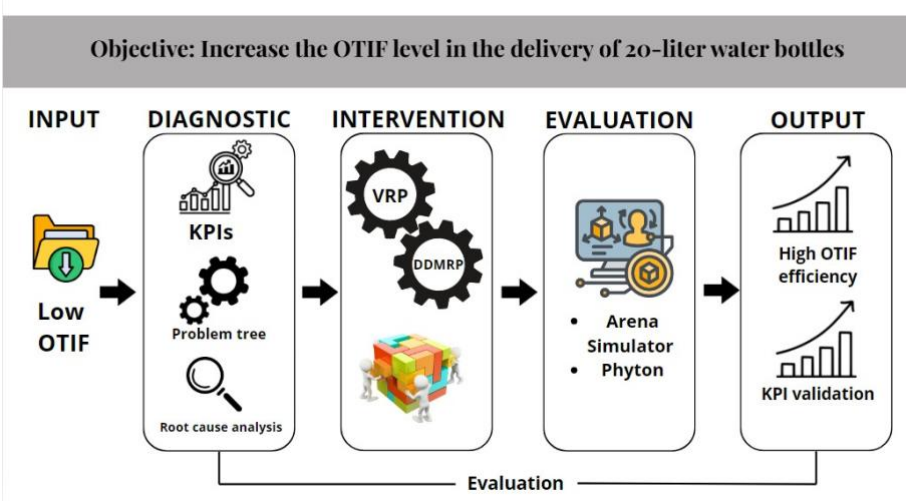


Figure 1: Model Components

Component 1 – Diagnosis: This stage establishes the baseline (As-Is) scenario through the analysis of operational indicators such as OTIF, incomplete deliveries, late deliveries, and total working time. Root-cause analysis techniques were applied to identify structural inefficiencies in route planning, workload allocation, and the lack of synchronization between inventory availability and dispatch operations. This component provides the analytical foundation for the subsequent intervention design.

Component 2 – Implementation: Based on the diagnostic findings, the improvement phase integrates two complementary tools. The Vehicle Routing Problem (VRP) restructures delivery routes using geolocation-based segmentation, demand grouping, vehicle-capacity constraints, and optimized distance matrices to reduce route irregularities and travel inefficiencies. In parallel, Demand-Driven MRP (DDMRP) stabilizes internal material flow through the configuration of dynamic inventory buffers aligned with consumption patterns. The integration of both tools ensures that routing decisions are supported by adequate product availability, generating a coherent To-Be operational scenario.

Component 3 - Validation: The effectiveness of the integrated model is evaluated through discrete-event simulation in Arena. Two simulation environments representing the As-Is and To-Be systems are compared using key performance indicators such as OTIF and operational time. Python supports customer-location processing and distance-matrix generation for model calibration. The comparative analysis quantifies

performance improvements and validates the impact of synchronizing routing optimization with dynamic inventory management.

2.7 Performance Indicators

To evaluate the impact of the proposed VRP–DDMRP–simulation model, a set of operational performance indicators was defined to quantify improvements in delivery performance, routing consistency, and production–distribution alignment.

On Time in Full (OTIF)

OTIF measures delivery punctuality and completeness, serving as the primary service-level indicator in distribution systems [4]

$$\text{OTIF} = \frac{\text{Total deliveries on time and complete}}{\text{Total deliveries}} * 100 \quad (1)$$

Late Deliveries (LD)

This indicator measures the proportion of deliveries performed outside the expected time window, often caused by route imbalance or inadequate vehicle assignment [8].

$$\text{LD} = \frac{\text{Number of late deliveries}}{\text{Total deliveries}} * 100 \quad (2)$$

Percentage of Incomplete Orders (PIO)

This indicator measures how many orders were not prepared correctly for dispatch, causing delays during loading. Only orders from the plant area are considered (not warehouse delays). [16].

$$\text{PIO} = \frac{\text{Number of orders not prepared}}{\text{Total deliveries}} * 100 \quad (3)$$

Percentage of Rescheduled Deliveries (PRD)

This indicator captures deliveries that must be rescheduled due to misalignment between production availability and dispatch sequencing. [17].

$$\text{PRD} = \frac{\text{Number of deliveries rescheduled due to coordination issues}}{\text{Total deliveries}} \quad (4)$$

3. Validation

3.1 Initial Diagnosis

The case study concerns a bottled water distribution company delivering 20-liter containers. The preliminary diagnosis revealed a significant service-level gap, reflected in low OTIF performance.

A total of 75.2% of non-compliances were related to late deliveries, mainly caused by inefficient route planning (40.0%), absence of structured order assignment (25.0%), and lack of geographic segmentation (10.2%). Additionally, 18.1% of the OTIF gap was associated with inventory-control deficiencies, including production–storage misalignment (12.2%) and unreliable inventory records (5.9%), leading to incomplete and rescheduled deliveries. The remaining 6.7% corresponded to minor operational causes. Figure 2 presents the problem tree linking these root causes to the proposed improvement tools.

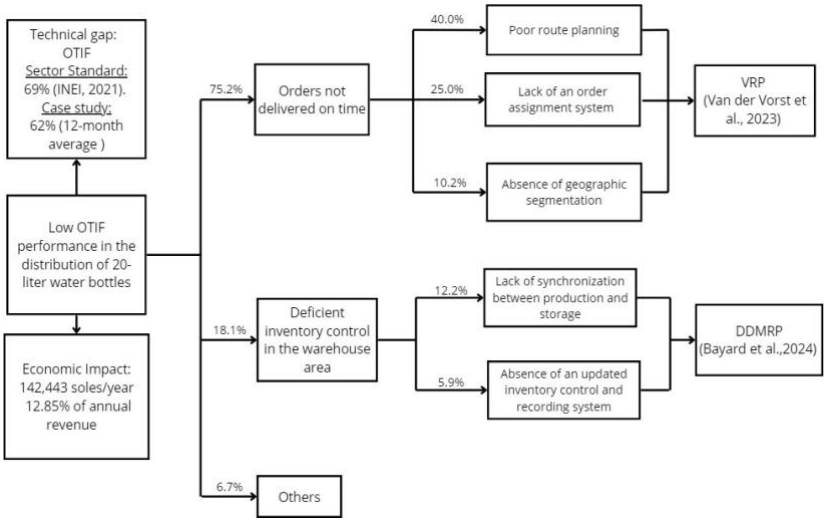


Figure 2: Problem tree

3.2 Implementation of the Pilot Plan

Stage 1 – Diagnostic segmentation and integrated data preparation:

The first stage consolidated and structured the operational data required for the integrated VRP–DDMRP model, including historical demand, customer locations, delivery frequencies, vehicle capacities, and consumption variability. Geographic coordinates were processed in Python to generate distance and travel-time matrices for routing analysis. Simultaneously, demand patterns, replenishment times, and inventory behavior were examined to support subsequent DDMRP buffer configuration. Customer segmentation into coherent delivery zones ensured consistent spatial clusters for VRP optimization while aligning demand groupings with inventory stability requirements. Figure 3 presents the distribution of delivery points and the baseline travel structure prior to optimization.

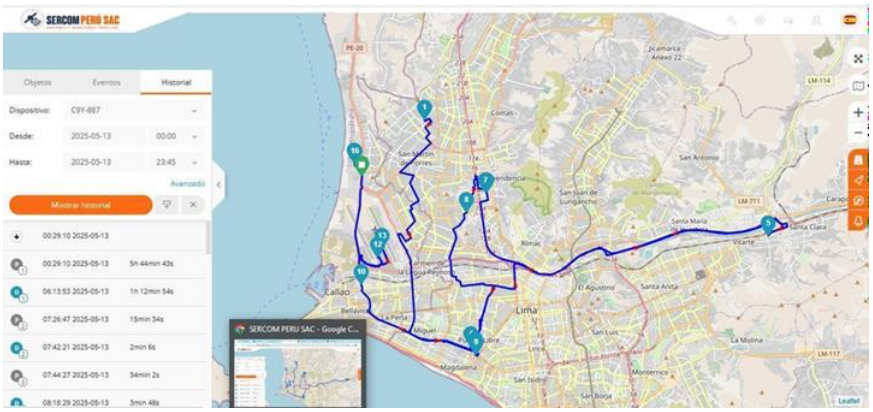


Figure 3: Delivery route prior optimization with VRP

Stage 2 – VRP optimization, route redesign and feasibility with inventory dynamics:

Following demand segmentation, the second stage applied the Vehicle Routing Problem using OR-Tools in Python to evaluate routing scenarios under vehicle-capacity, demand, distance, and service-time constraints. The algorithm generated optimized visitation sequences, whose feasibility was validated against inventory profiles derived from the DDMRP demand analysis. This ensured alignment between routing plans and replenishment capacity, reducing incomplete orders and rescheduling risks. The optimized routes replaced manually assigned paths, improving delivery consistency while maintaining coherence with material-flow conditions. Figure 4 presents the optimized routing configuration, illustrating reductions in detours, route overlap, and sequencing inefficiencies.

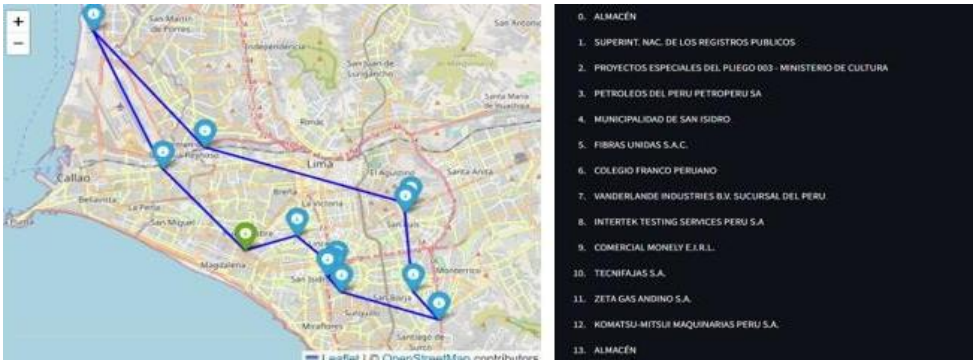


Figure 4: Map with optimized routes after VRP application

Stage 3 – DDMRP buffer configuration and synchronization with optimized distribution:

The third stage involved designing and calibrating DDMRP decoupling buffers to stabilize product availability and synchronize internal material flows with the optimized routing structure obtained from the VRP model. Products were classified according to demand variability, lead-time characteristics, and average consumption, enabling the definition of strategic buffer positions within the warehouse. The red, yellow, and green zones were calculated using consumption rates and replenishment times, allowing operators to visually prioritize inventory decisions while ensuring sufficient stock to support the daily routing sequences. Table 1 summarizes the assigned buffer values for each zone, detailing the calculated inventory levels used to dimension the decoupling points. By aligning buffer levels with the optimized distribution plan, the configuration significantly reduced incomplete deliveries, last-minute adjustments, and reprogrammed orders, reinforcing the coherence and reliability of the integrated improvement model. Figure 5 shows the resulting buffer model applied to the bottled-water inventory.

Table 1: Assigned Inventory Levels for DDMRP Decoupling Buffers

Zones	Value (20-liter containers)
Red Zone	0 - 273
Yellow Zone	273 - 427
Green Zone	427 - 700

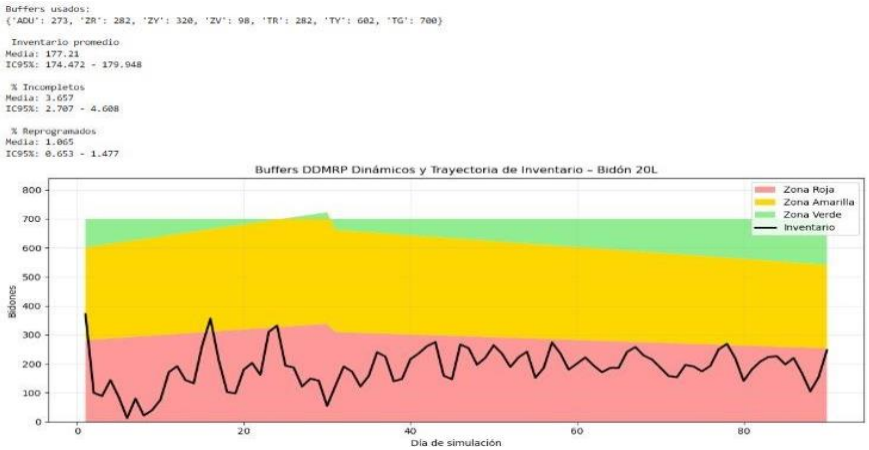


Figure 5: DDMRP Buffers and Inventory Trajectory for 20-Liter Containers

3.3 Simulation – Based Validation

The distribution process was replicated using discrete-event simulation to compare the baseline (As-Is) and improved (To-Be) scenarios. The As-Is model reproduces the current operating conditions characterized by

manual routing and inventory–dispatch misalignment. Figure 6 shows the simulation model of the current distribution process.

Current situation: distribution process

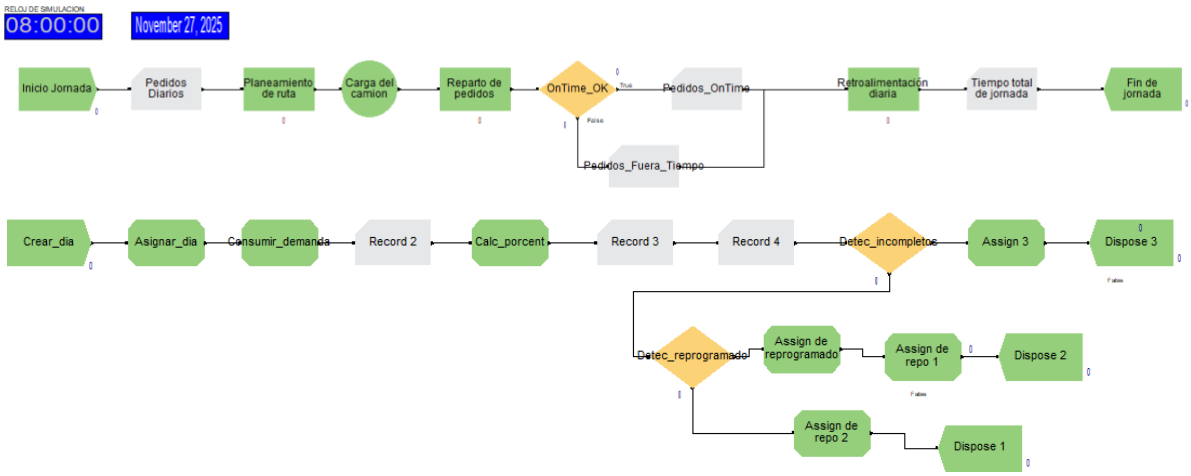


Figure 6: Simulation model of the current distribution process

The To-Be model incorporates VRP-based route optimization and DDMRP-based buffer stabilization to synchronize routing and inventory availability. Figure 7 presents the simulation model of the improved distribution process.

Improved situation: distribution process

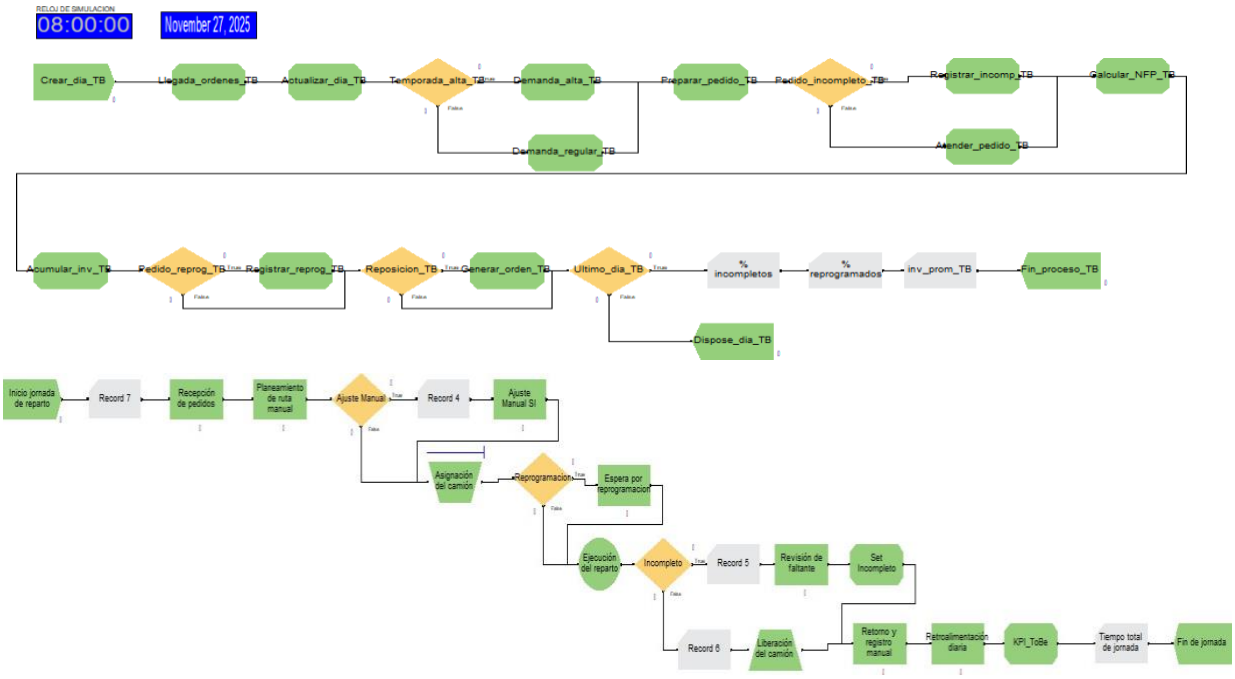


Figure 7: Simulation model of the improved distribution process

The operational performance of both scenarios was evaluated using the defined key indicators. Table 2 summarizes the comparative results, where the “Variability” column represents the relative change between the As-Is and To-Be scenarios.

Table 2: Comparative table of results

Indicator	AS IS	TO BE	Variability
On Time In Full (OTIF)	62.0%	69.0%	7.0%
Late delivery	40.0%	18.6%	21.4%
Incomplete orders	12.1%	4.0%	8.1%
Rescheduled deliveries (RD)	5.9%	1.5%	4.4%

4. Economic and financial evaluation

The economic and financial feasibility of the proposed model was evaluated through a five-year cash flow projection using a cost of capital (COK) of 12%. The analysis incorporated the quantified operational improvements derived from the simulation results and the associated implementation costs.

The evaluation yielded a Net Present Value (NPV) of S/ 238,981.66, confirming value generation over the project horizon. The Internal Rate of Return (IRR) significantly exceeded the discount rate, indicating strong profitability relative to the assumed opportunity cost of capital.

Overall, the results support the financial viability of implementing the integrated VRP–DDMRP model, demonstrating that operational efficiency improvements translate into measurable economic benefits.

5. Results and discussion

The validation results confirm that integrating VRP improved delivery punctuality, reducing late deliveries to 18.6%, consistent with reported routing gains of 5–8% and efficiency improvements up to 14% [4]–[6], [9]. The stronger impact observed is linked to the initial context of manual routing and high variability. Simultaneously, DDMRP stabilized inventory performance, achieving 3.7% incomplete orders and 1.1% rescheduled deliveries, aligning with prior evidence on stockout reduction and variability control through dynamic buffers [15], [18]. Overall, synchronizing routing optimization with demand-driven inventory control enhanced service reliability in this reactive distribution setting.

However, limitations remain. The analysis relied on one year of historical data, limiting seasonal assessment. The VRP model excluded traffic conditions and time-window constraints, and DDMRP buffers required manual calibration due to limited variability data. Validation was conducted exclusively through simulation, restricting direct generalization. Future research should incorporate real-time traffic data, heterogeneous fleets, predictive buffer parametrization, and broader product portfolios to strengthen robustness and scalability.

6. Conclusion and outlook

This study demonstrated that integrating VRP and DDMRP within a unified decision-support framework effectively mitigates the misalignment between routing decisions and inventory availability identified during diagnosis. The proposed model increased OTIF from 62% to 69% while reducing incomplete and late deliveries, confirming the complementary impact of route optimization and dynamic buffer stabilization. The intervention also proved financially viable, generating a positive Net Present Value (NPV) of S/ 238,981. By synchronizing routing efficiency with inventory stability, the study contributes empirical evidence to the limited literature integrating distribution optimization and demand-driven planning in SME contexts.

Future research should advance this framework through pilot implementation under real traffic conditions, dynamic buffer parametrization using predictive techniques, and validation across broader product portfolios. These extensions would strengthen robustness, scalability, and long-term applicability in real operational environments.

Appendix 1: Comparative table of the components of the Proposal Model vs. Literature Review

	Vehicle Routing with Time Windows Applying the Clarke–Wright Algorithm for a Medical Supplies Company		A Vehicle Routing Problem with a Time-Window Approach to Improve the Delivery Process		Multi-objective Vehicle Routing Problem with Simultaneous Deliveries and Pickups and Emission Minimization		Effective production planning for purchased part under long lead time and uncertain demand: MRP vs demand-driven MRP		Simulation of an Inventory Policy Based on the Demand-Driven MRP Methodology Using a Petri Net Approach		Optimizing Last-Mile Distribution in an Emerging-Economy Bottled Water Firm Through VRP and DDMRP Integration	
	Barcia, K. F., Gutiérrez, R. A., & González, V. H.		U,Pineda & H,Carabalí		J. Ruiz-Meza		M,Shofa, A,Moeis & N,Restiana		A,Paredes, K.Ciro & J,Jaramillo		Castañeda-Ríos., Zamalloa-Peralta, V., Ruiz-Ruiz, M., Quiroz-Flores, J	
Component 0: Problem diagnosis - Literature Review – Engineering Tool: VRP diagnostic mapping – Case Study: Bottled water distribution SME – Metrics: OTIF, incomplete deliveries, late deliveries, excess travel distance	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓
	Engineering Tool	✓	Engineering Tool		Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓
	Case Study	✓	Case Study		Case Study		Case Study		Case Study	✓	Case Study	✓
	Metric	✓	Metric	✓	Metric	✓	Metric	✓	Metric	✓	Metric	✓
Component 1: Implementation - Literature Review – Engineering Tool: VRP optimization model + DDMRP buffers – Case Study: Bottled water distribution SME – Metrics: OTIF improvement, route efficiency, buffer performance, stock availability	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓
	Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓
	Case Study		Case Study		Case Study		Case Study		Case Study		Case Study	✓
	Metric	✓	Metric	✓	Metric		Metric	✓	Metric	✓	Metric	✓
Component 2: Validation - Literature Review – Engineering Tool: Simulation – Case Study: Bottled water distribution SME – Metrics: OTIF, kilometers traveled, delivery time, inventory stability, DDMRP buffer audit	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review		Literary Review	✓
	Engineering Tool		Engineering Tool	✓	Engineering Tool		Engineering Tool	✓	Engineering Tool	✓	Engineering Tool	✓
	Case Study	✓	Case Study		Case Study		Case Study	✓	Case Study		Case Study	✓
	Metric	✓	Metric	✓	Metric	✓	Metric	✓	Metric	✓	Metric	✓
	Total	10	Total	8	Total	7	Total	9	Total	9	Total	12

References

[1] Ministerio de la Producción (PRODUCE), 2022. Micro, Small and Medium Enterprises in Figures 2022. Lima, Peru.

[2] Euromonitor International, 2023. Bottled Water in Peru: Industry Overview and Forecasts. Passport Database.

[3] COMEXPERU, 2023. Annual Report 2023. Lima, Peru.

[4] Ruiz-Meza, J., 2023. Multi-objective vehicle routing problem with simultaneous deliveries and pickups and emission minimization. Journal of Applied Research and Technology 21 (3), 435–448. https://doi.org/10.4067/S0718-33052021000300435

[5] Barcia, K., Gutierrez, R., Gonzales, V., 2020. Vehicle routing with time windows applying the Clarke–Wright algorithm for a medical supplies company. Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology. https://doi.org/10.18687/LACCEI2020.1.1.110

[6] Rodriguez-Vasquez, W., 2020. Modeling a vehicle routing problem with multiple depots, time windows, and a heterogeneous fleet for a courier service. Journal of Applied Research and Technology 18 (1), 1–12. https://doi.org/10.4067/S0718-07642020000100207

[7] Inga, A., Rojas, J., Basurto, M., Callupe, S., Luis, K., Morales, Y., 2020. Improvement of the distribution system of a cement and construction materials trading company through the use of mathematical modeling. Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology. https://doi.org/10.18687/LACCEI2020.1.1.180

[8] Pineda, U., Carabali, H., 2020. A vehicle routing problem with a time-window approach to improve the delivery process. Ingeniería e Investigación 40 (2), 23–31. https://doi.org/10.14483/23448393.15271

[9] Shofa, M., Moeis, A., Restiana, N., 2018. Effective production planning for purchased parts under long lead time and uncertain demand: MRP vs demand-driven MRP. IOP Conference Series: Materials Science and Engineering 337 (1), 012055. https://doi.org/10.1088/1757-899X/337/1/012055

[10] Gonzales-Leon, L., Salcedo-Portocarrero, N., Quiroz-Flores, J., 2023. Inventory management model for reducing stockout rate by applying lean warehousing and DDMRP tools in SMEs in the commercial sector. World Conference on Systems Engineering Proceedings. https://doi.org/10.18178/wcse.2023.06.055

[11] Tilk, C., Olkis, D., Irnich, S., 2021. The last-mile vehicle routing problem with delivery options. OR Spectrum 43 (2), 321–356. https://doi.org/10.1007/s00291-021-00633-0

[12] Misopoulos, F., 2020. Carbon emissions of bottled water sector supply chains: a multiple case-study approach. International Journal of Logistics Research and Applications 23 (2), 178–194. https://doi.org/10.1080/13675567.2019.1626815

[13] Soto-Concha, D., Galvez, D., Soto, R., Canca, D., 2025. The vehicle routing problem with satellites utilization: a systematic review of the literature. Mathematics 13 (7), 1092. https://doi.org/10.3390/math13071092

[14] Van der Vorst, J., Tromp, S., Van der Zee, D., 2023. Integral logistics solutions for retail chains: synchronization of supply and demand. International Journal of Physical Distribution and Logistics Management 53 (4), 389–410. https://doi.org/10.1108/IJPDLM-01-2022-0016

[15] Miclo, R., Dolgui, A., Hnaien, F., 2020. Demand driven MRP: review and opportunities for future research. International Journal of Production Research 58 (6), 1649–1673. https://doi.org/10.1080/09537287.2019.1671747

[16] Dolgui, A., Ivanov, D., Sokolov, B., 2023. Demand driven MRP in dynamic and uncertain environments: a systematic review. International Journal of Production Research 61 (7), 2156–2174. https://doi.org/10.1080/00207543.2021.2007802

[17] Dessevre, G., Benali, M., 2024. Analysis of a dynamic capacity management approach in DDMRP: application on a real industrial case. Journal of Industrial Engineering and Management 17 (4), 742–760. https://doi.org/10.3926/jiem.6045

[18] Paredes, A., Ciro, K., Jaramillo, J., 2022. Simulation of an inventory policy based on the demand-driven MRP methodology using a Petri net approach. Ingeniería e Investigación 42 (1), 1–10. https://doi.org/10.14483/23448393.18002

[19] Reis, J., 2023. Exploring applications and practical examples by streamlining material requirements planning (MRP) with Python. Logistics 7 (4), 91. https://doi.org/10.3390/logistics7040091

[20] Liu, Y., Zhang, L., Wang, S., 2023. Mapping water scarcity risks in global supply chain networks: a multi-sector input–output approach. International Journal of Logistics Research and Applications 27 (3), 321–339. https://doi.org/10.1080/13675567.2024.2330591

Biography

Boris Castañeda Rios is an Industrial Engineering student at the University of Lima, where he has been enrolled since 2020 and ranks in the upper third of the Faculty of Engineering. He maintains a strong professional focus on supply chain management and has developed particular interest in areas such as inventory optimization, operations planning, and process improvement. He aims to build a solid career in industries that value operational efficiency especially within sectors like manufacturing and consumer goods.

Veralucia Zamalloa Peralta is an Industrial Engineering student at the University of Lima and a member of the top third of the School of Engineering. Her academic interests focus on operations and supply chain management, particularly demand planning, commercial management, and logistics optimization. She aims to develop a professional career in dynamic, consumer-oriented sectors, contributing analytical and innovation-driven solutions to improve operational performance.

Marcos Fernando Ruiz-Ruiz holds a PhD in Education Sciences from the Pontifical Catholic University of Peru. He is an Industrial Engineer from the University of Lima and holds a master's degree in education from the University of Piura. He is a Renacyt-accredited researcher and completed an Expert Diploma in Future Studies at UNED (Spain). His research focuses on structural analysis, foresight, and gender approaches within STEM education.

Juan Carlos Quiroz-Flores holds a PhD in Business Administration from the Universidad Nacional Mayor de San Marcos. He is an Industrial Engineer from the University of Lima, an MBA graduate from ESAN University, and a Black Belt in Lean Six Sigma from the Peruvian University of Applied Sciences. He is a professor and research advisor at the University of Lima. His professional background spans more than 25 years in supply chain management, lean operations, process improvement, and productivity enhancement.

8th Conference on Production Systems and Logistics

Data-Driven Lean Warehousing And Hybrid Simulation To Improve Perfect Order Rate In Medical Supply SMEs

Carlos Daza-Llanos, Said Rojas-Lavado, Marcos Fernando Ruiz-Ruiz, and Juan Carlos Quiroz-Flores

Carrera de Ingeniería Industrial, Universidad de Lima, Lima, Perú.

Abstract

Medical supply SMEs in emerging markets face persistent challenges in warehouse efficiency, inventory accuracy, and service reliability, as noted in prior logistics research. This study addresses these constraints by targeting the low Perfect Order Rate (POR) of a Peruvian medical supply importer, affected by inventory discrepancies, picking congestion, and documentation errors. A data-driven Lean Warehousing model was developed, integrating 5S, ABC-based slotting, cycle counting, and poka-yoke mechanisms, and validated through a hybrid method combining a six-week pilot implementation with discrete-event simulation in Arena and FlexSim. Results show improvements in On-Time Delivery from 91.35% to 94.15%, Defect-Free Products from 95.12% to 98.89%, and Process Cycle Efficiency from 5% to 37%, raising the projected POR from 82.51% to 88.30%. These findings highlight the operational and socioeconomic relevance of Lean–Digital practices for SMEs and encourage further exploration of data-driven enhancements in regulated supply chains.

Keywords

Lean Warehousing; Perfect Order Rate; Discrete-Event Simulation; Medical Supply Chains; Operational Reliability.

1. Introduction

The medical supply sector plays a critical role in healthcare delivery across Latin America, where rising demand and strong dependence on imported products have intensified operational pressures on small and medium-sized enterprises (SMEs) [1], [2]. These firms frequently operate with limited automation, fragmented warehouse layouts, and constrained resources, conditions that reduce logistics reliability and amplify the impact of disruptions throughout the supply chain [3]. In the company analyzed, several deficiencies were linked to declining warehouse performance, including inaccurate inventory records, congestion in picking zones, disorganized SKU placement, and insufficient standardization of operating procedures. These issues directly affected the firm's core logistics indicators—On-Time Delivery (OTD), Defect-Free Products (DPP), Process Cycle Efficiency (PCE), and Perfect Order Rate (POR)—with POR reaching only 82.51%, well below the levels reported in mature supply chains [4]. This situation defines the problem statement: the need to increase POR by improving inventory accuracy, picking efficiency, and dispatch reliability within a resource-limited SME importer.

DOI: https://doi.org/10.15488/21025
ISSN: 2701-6277

Previous studies show that Lean Warehousing tools such as 5S, slotting, cycle counting, and poka-yoke have proven effective in reducing errors, eliminating waste, and enhancing warehouse flow [5]. In parallel, discrete-event simulation has been widely employed to test operational changes and evaluate layout redesigns in controlled virtual environments, enabling evidence-based decision-making without interrupting day-to-day operations [5]. However, existing research rarely integrates Lean Warehousing with data-driven digital support and hybrid simulation, and almost no work has evaluated such integration in medical supply SMEs operating in emerging markets. This lack of unified, empirically validated models constitutes the research gap addressed in this study.

To respond to this gap, the study develops a data-driven Lean Warehousing model that incorporates 5S, ABC-based slotting, cycle counting, and poka-yoke mechanisms, supported by digital data capture and hybrid validation (six-week pilot + Arena + Flex Sim). The novelty of this research lies in demonstrating how a coherent Lean–Digital framework can enhance POR in medical supply SMEs, and in providing empirical and simulation-based evidence of its effectiveness in a highly regulated and resource-constrained environment.

2. State of the art

2.1 Perfect Order Rate (POR)

Medical-device supply chains face critical structural and operational challenges that directly compromise key performance indicators such as Perfect Order Rate (POR), On-Time Delivery (OTD), and overall logistics costs. The sector's low adoption of automation-currently at 14.9%, significantly below the industrial average-intensifies exposure to human error and process inefficiencies, limiting operational reliability [6]. Research identifies inherent fragility in healthcare supply chains, driven by rigid organizational structures and complex regulatory requirements that result in elevated operating costs and reduced responsiveness to demand fluctuations [7], [8]. Furthermore, dependence on single-source suppliers heightens operational risk and exposure to cost overruns during disruptions [9], while an excessive reliance on intermediaries and infrastructure limitations inflates distribution costs and obstructs process standardization [6]. Managing large product portfolios demands advanced classification systems and layout redesigns to handle variety efficiently [6], [10]; studies indicate that even organizations with fewer SKUs suffer significant productivity losses due to poor spatial organization [10]. Finally, regulatory diversity lengthens operational cycles, delaying market access [8], and although intelligent technologies such as AI and blockchain offer viable solutions to reduce operational time and costs, their adoption remains constrained among medium and small distributors [6], [11].

2.2 Common Challenges in Medical Importers and Distributors

Recent literature indicates that medical importers and distributors operate in an environment increasingly constrained by regulatory complexity and operational bottlenecks. A primary challenge is the regulatory barrier to market entry. The medical device industry faces significant hurdles due to evolving global regulations, which directly impact innovation cycles and restrict market access for new technologies [7]. This regulatory landscape is further complicated by a lack of standardization across regions; the absence of global regulatory harmonization compels distributors to duplicate validation practices, creating administrative inefficiencies that delay the supply of critical pharmaceutical and medical products [9].

Beyond regulatory compliance, operational efficiency in order fulfillment remains a critical pain point. While e-commerce has advanced rapidly in other sectors, adapting these models to high-stakes environments proves difficult. Developing efficient fulfillment models that balance speed with the high accuracy required in logistics presents a significant challenge, particularly in sectors where error tolerance is minimal [10].

Furthermore, the healthcare logistics sector exhibits a notable lag in technological adoption compared to other industries. Exploratory research reveals that the application of logistics automation in the healthcare industry is still in a nascent stage, hindered by technical barriers and investment hesitancy [11]. This slow digital transformation limits supply chain visibility; the insufficient adoption of Industry 4.0 technologies—such as RFID and decision support systems—prevents firms from achieving real-time traceability and optimizing performance against demand fluctuations [13].

2.3 Lean Improvement Tools for Warehouse Optimization

Optimizing the Perfect Order Rate (POR) demands the deployment of an integrated toolkit designed to enforce operational discipline, enhance traceability, and systematically eradicate errors. The 5S methodology serves as the foundational framework for this transformation; empirical evidence confirms its efficacy in recovering usable workspace and drastically reducing retrieval times when sustained through rigorous audits and continuous training [14], [5]. Furthermore, strategic slotting and warehouse redesign are critical for optimizing internal material flows and volumetric utilization, a priority in medical-device environments where strict lot traceability and labeling protocols dictate both productivity and regulatory compliance [15], [6]. To safeguard inventory integrity, Cycle Counting strategies have proven superior to traditional methods, significantly mitigating Inventory Record Inaccuracy (IRI) and enhancing discrepancy detection capabilities in dynamic warehouse settings [16]. Concurrently, the implementation of robust error-proofing mechanisms (*poka-yoke*), such as standardized kitting and double-verification systems, directly reduces fulfillment failures and reverse logistics costs [17]. Given the sector's constrained automation landscape, visual management systems and scalable lightweight technologies offer high-impact, accessible solutions for small and medium-sized enterprises [11].

2.4 Discrete Event Simulation for Layout Optimization

While Lean tools address operational discipline, structural changes require rigorous validation prior to implementation to mitigate risks. Discrete Event Simulation (DES) has proven to be a critical methodology for validating warehouse redesigns and picking strategies without disrupting ongoing operations. Recent studies utilizing simulation software (e.g., Arena, AnyLogic) have demonstrated the ability to achieve significant performance gains, such as a 36.94% reduction in travel distances and improvements in order fulfillment efficiency ranging from 8.33% to 13.20% depending on the strategy employed [5], [9]. Furthermore, integrated models combining storage location assignment (Slotting) with picker routing allow for the precise evaluation of spatial constraints and operator movements, ensuring that theoretical improvements in travel distance translate into actual productivity gains and optimized material handling [18], [15]. This predictive capability makes simulation and mathematical validation essential complements to Lean implementations in complex logistics environments.

While prior studies highlight the individual benefits of Lean tools and simulation-based analysis, the literature provides limited evidence of integrated, data-driven Lean Warehousing models validated through hybrid simulation in medical supply SMEs, reinforcing the need for the approach developed in this study.

3. Proposed Model

A comprehensive state-of-the-art analysis revealed several thematic gaps in existing models regarding logistics optimization in emerging markets. For each component listed in **Appendix 1**, we established specific evaluation criteria based on literature review, engineering tool application, and performance metrics. We searched for models that applied specific techniques in stocking contexts and measured indicators such as service level, picking time, and inventory accuracy.

The analysis showed that most existing models met only partial criteria, focusing either solely on soft tools (like 5S) or theoretical simulation, without integrating a holistic approach. Due to limited research on Lean Warehousing specifically in the medical supply import sector, we used models from similar distribution

industries as references. To address these gaps, the proposed model integrates successful elements from previous research but is specifically designed for SMEs with limited resources. It incorporates ABC analysis to improve traceability, 5S to enhance organization, Cycle Counting for inventory precision, and Poka-Yoke for error reduction.

3.1 Structure of the Data-Driven Model

The study identifies a gap in Lean Warehousing implementation within the medical supply import sector, limiting efficient management of critical inventory. To address this, a data-driven model is proposed to enhance operational reliability and competitiveness. The objective is to increase the Perfect Order Rate (POR) by reducing picking times, eliminating dispatch errors, and ensuring inventory accuracy.

As shown in **Figure 1**, the process begins with a quantitative diagnosis using Value Stream Mapping and Pareto analysis on 12 months of historical data. Then, tools such as 5S, ABC-based Cycle Counting, Slotting, and Poka-Yoke are implemented to optimize resources and flow. Finally, the model is validated through a hybrid approach combining physical pilots and discrete-event simulation (Arena and FlexSim). This dual validation uses statistically calibrated inputs (Goodness-of-Fit tests, p-value > 0.05) to project the sustainability of improvements beyond the pilot phase.

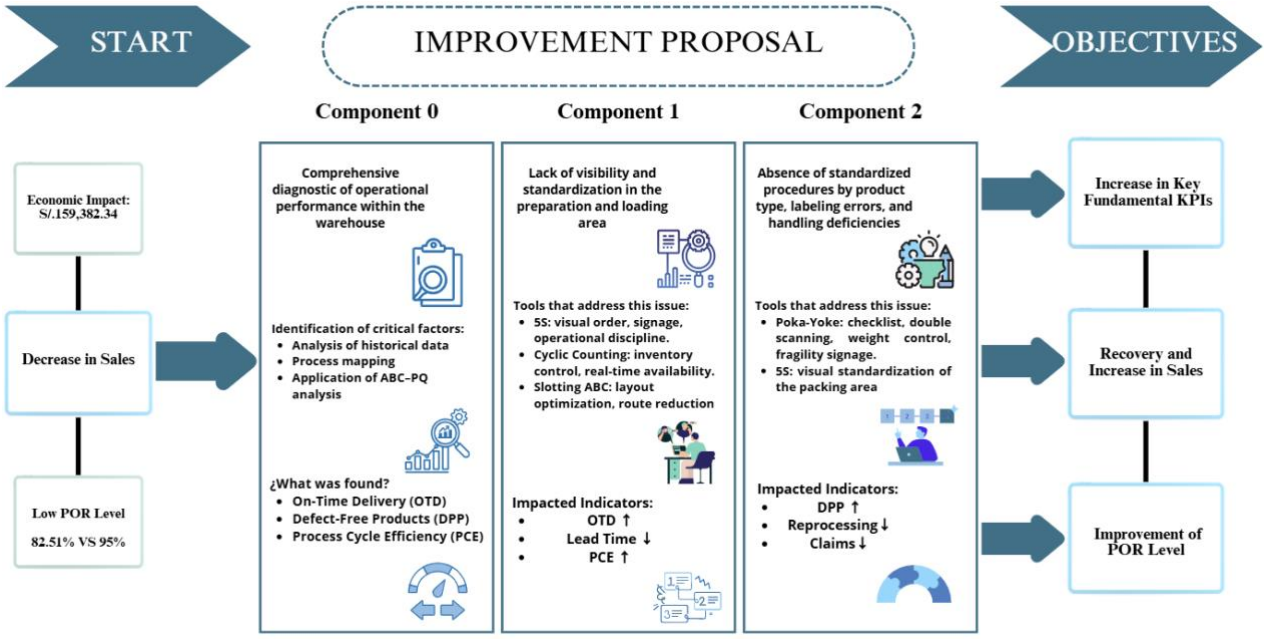


Figure 1: Model Components

The model is structured into three sequential components that ensure logical progression from diagnosis to standardization and optimization:

Component 0: Baseline Performance Assessment & Root-Cause Diagnosis This component establishes the empirical and analytical foundation of the improvement proposal. Through the examination of 12 months of historical fulfillment data, process mapping, and the application of Data-Driven ABC–PQ analysis, the critical drivers behind the low Perfect Order Rate (POR) are identified. The assessment evaluates deviations in the time, quantity, condition, and documentation dimensions, revealing gaps in On-Time Delivery (OTD), Defect-Free Products (DPP), and Process Cycle Efficiency (PCE). This diagnosis provides a clear understanding of the operational constraints and serves as the statistical baseline against which the improvement components are strategically designed.

Component 1: Operational Synchronization & Flow Velocity This component transforms the warehouse environment by enforcing strict operational discipline to guarantee flow visibility. By fusing 5S standards with ABC-based Cycle Counting, it ensures real-time inventory reliability. Simultaneously, strategic Slotting, derived from picking frequency analysis, is deployed to minimize travel waste. These interventions directly maximize On-Time Delivery (OTD) and Inventory Accuracy, while significantly boosting Process Cycle Efficiency (PCE), collectively strengthening the time and quantity dimensions of the POR.

Component 2: Zero-Defect Fulfilment & Quality Assurance Designed to safeguard the integrity of critical medical supplies, this component establishes a "zero-error" barrier in the packing phase. Through the deployment of rigorous Poka-Yoke mechanisms, including double-scan verification and physical templates, it drastically reduces human variability. This approach ensures a substantial rise in the Defect-Free Products (DPP) rate, directly securing the quality and documentation dimensions of the POR.

3.2 Model Indicators

To quantify the impact of the proposed Lean Warehousing model on operational performance, four key indicators were defined based on the initial diagnosis and industry standards:

On-Time Delivery (OTD)

Research emphasizes that optimizing this indicator is crucial to avoiding bottlenecks in distribution and ensuring customer satisfaction [4].

$$OTD = \frac{Orders\ Delivery\ On\ Time}{Total\ Orders} * 100 \quad (1)$$

Defect-Free Products (DPP)

This indicator evaluates the quality of the dispatch process by tracking orders free from damage or packaging errors, directly linked to error-proofing strategies **[26].**

$$DPP = \frac{Orders\ Without\ Damage}{Total\ Orders} * 100 \quad (2)$$

Process Cycle Efficiency (PCE)

This metric measures the workflow efficiency by comparing value-added time against the total lead time, essential for validating Lean implementations in SMEs [19].

$$PCE = \frac{Value\ Added\ Time}{Total\ Lead\ Time} * 100 \quad (3)$$

Perfect Order Rate (POR)

The global metric for logistics reliability, defined as the product of on-time, complete, undamaged, and correctly documented orders [4].

$$POR = (\%OTD) * (\%Complete) * (\%DPP) * (\%\ Documentación) \quad (4)$$

4. Validation

To validate the proposed data-driven Lean Warehousing model, a hybrid approach was used, combining a six-week pilot implementation with discrete-event simulation. Empirical data from the pilot—reductions in picking cycle time from 5S and slotting, along with near-elimination of documentation errors through poka-yoke—were used to calibrate the To-Be scenario in Arena and Flex Sim. The As-Is simulation was built using diagnostic measurements, while the To-Be model incorporated pilot-derived performance improvements. Confidence intervals for both scenarios were compared, and paired t-tests were applied to assess statistical significance across key indicators. This hybrid validation strategy provided rigorous

evidence that the integrated Lean–Digital framework generated measurable improvements in POR and its component indicators.

4.1 Initial Diagnosis

The case study company specializes in the import and commercialization of medical laboratory supplies, serving hospitals and clinics nationwide. The preliminary diagnosis reveals a technical gap in the warehouse's reliability, with a Perfect Order Rate (POR) of 82.51%, which is below the expected industry standard of 95%. This deficiency generates an economic impact equivalent to 6.54% of annual sales due to penalties and rework.

This performance gap is mainly due to late deliveries (46.4%), damaged products (26.19%), and incomplete orders (20.24%). During diagnostic data collection, several root causes were identified: the absence of standardization and visibility in the preparation flow (affecting PCE), concentration of outflows exceeding loading zone capacity (affecting OTD), non-standardized packing without technical validation, and a layout not adjusted to product rotation and fragility (affecting DPP). **Figure 2** illustrates the problem tree developed during the diagnostic phase, mapping the identified root causes to the associated reasons and the central performance gap.

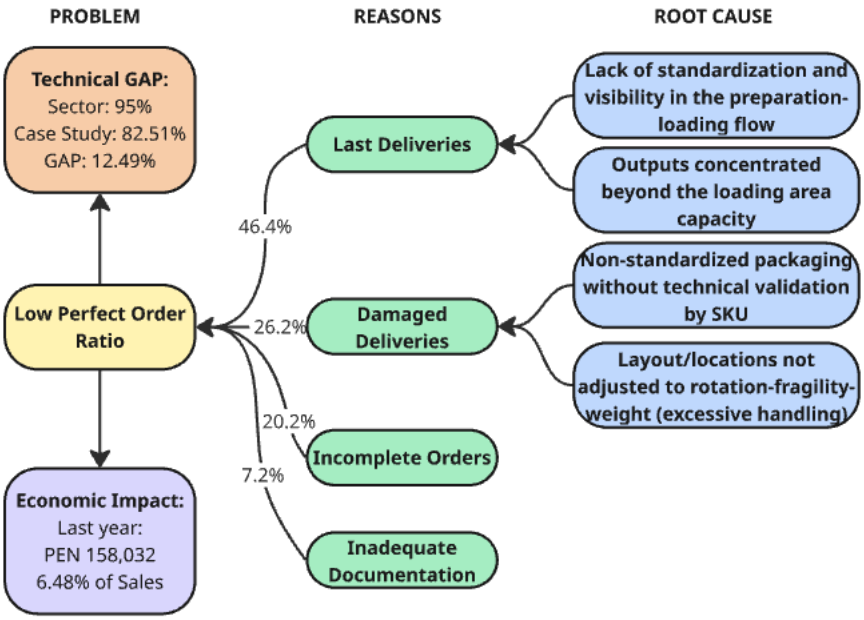


Figure 2: Problem tree

4.2 Implementation of the Pilot Plan

4.2.1 Stage 1 – 5S & Visual Management

The implementation began with a rigorous Seiri (Sort) phase, applying red-tagging campaigns to identify and remove obsolete items, which resulted in the recovery of 32% of usable operational space. In the Seiton (Set in Order) phase, visual standards were applied to demarcate aisles and staging zones according to NTP 399.010 safety standards. Weekly audits confirmed that the organization of critical areas improved from 40% to 70%, while the average search time per SKU dropped drastically from 4.5 to 2.8 minutes. These improvements were reflected visually in the warehouse layout, as shown in **Figure 3**, which illustrates the contrast between the initial disorganized state and the conditions achieved after the 5S deployment.

Initial situation (Before)	Status before 5S (After)

Figure 3: Warehouse before and after 5S deployment

4.2.2 Stage 2 – Cycle Counting (ABC):

To replace annual physical inventories, a continuous counting routine was established. The protocol involved daily verification of Class A items (21 critical SKUs representing 75% of flow) by locking locations during counting to ensure data integrity. The process involved two operators (counter and recorder) to minimize bias. The pilot results showed an increase in Inventory Record Accuracy (IRA) from 85% to 94%. Furthermore, the time required per count decreased from 4.5 to 2.9 minutes, proving the efficiency of the new method. These improvements in accuracy and counting efficiency were also reflected in the physical reorganization of the storage zones, as shown in **Figure 4**, which illustrates the before-and-after conditions of the ABC-classified areas and the operators performing the cycle counting procedure.

Initial situation (Before)	Final Situation (After)

Figure 4: ABC classified storage zones and operator performing cycle counting with barcode scanner

4.2.3 Stage 3 – Poka-Yoke & Digital Control

Error-proofing mechanisms were integrated into the dispatch process using a double-check system. As shown in **Figure 5**, visual standards were deployed to guide the packing sequence (steps 1-9). Additionally, a digital log via Power Apps combined with handheld scanners was introduced to validate SKUs and batches in real-time. This intervention reduced the rate of returns due to damage or incorrect labelling from 3.1% to 1.6%.

Figure 5: Implementation of Visual Packing Standards and Digital Poka-Yoke using Power Apps.

In addition to the qualitative improvements illustrated in Figure 5, the quantitative effects of each pilot intervention are summarized in **Table 1**. This table consolidates the measurable outcomes of 5S, Cycle Counting, and Digital Poka-Yoke across the three stages, providing a clear overview of their impact on search time, inventory accuracy, and return rates.

Table 1: Comparative table of results of the pilot test

Tool	Indicator	As-Is	Pilot Result	Impact
5S	Search Time (min/SKU)	4.5	2.8	-37.70%
Cycle Counting	Inventory Accuracy (IRA)	85%	94%	+9 p.p.
Poka-Yoke	Return Rate (Damage/Labeling)	3.10%	1.60%	-48.40%

4.3 Simulation and Validation

Following the operational pilots, the integrated model was validated using a dual simulation approach to ensure both statistical robustness and spatial efficiency.

Statistical Validation (Arena)

The complete operational flow was simulated using Arena Software (30 replications) to verify the impact on process indicators. The initial model ("As-Is") confirmed a low efficiency (PCE ~5%) due to bottlenecks in picking and dispatch. The improved "To-Be" model incorporated the parameters obtained from the pilots (94% inventory accuracy, reduced search times). The results revealed a statistically significant improvement ($p < 0.05$) in all key metrics: On-Time Delivery (OTD) increased to 94.15%, and Defect-Free Products (DPP) rose to 98.89%. The resulting operational flow was represented through the Arena simulation model, as shown in **Figure 6**, which depicts the sequence of warehouse activities and the key performance indicators used for the statistical validation.

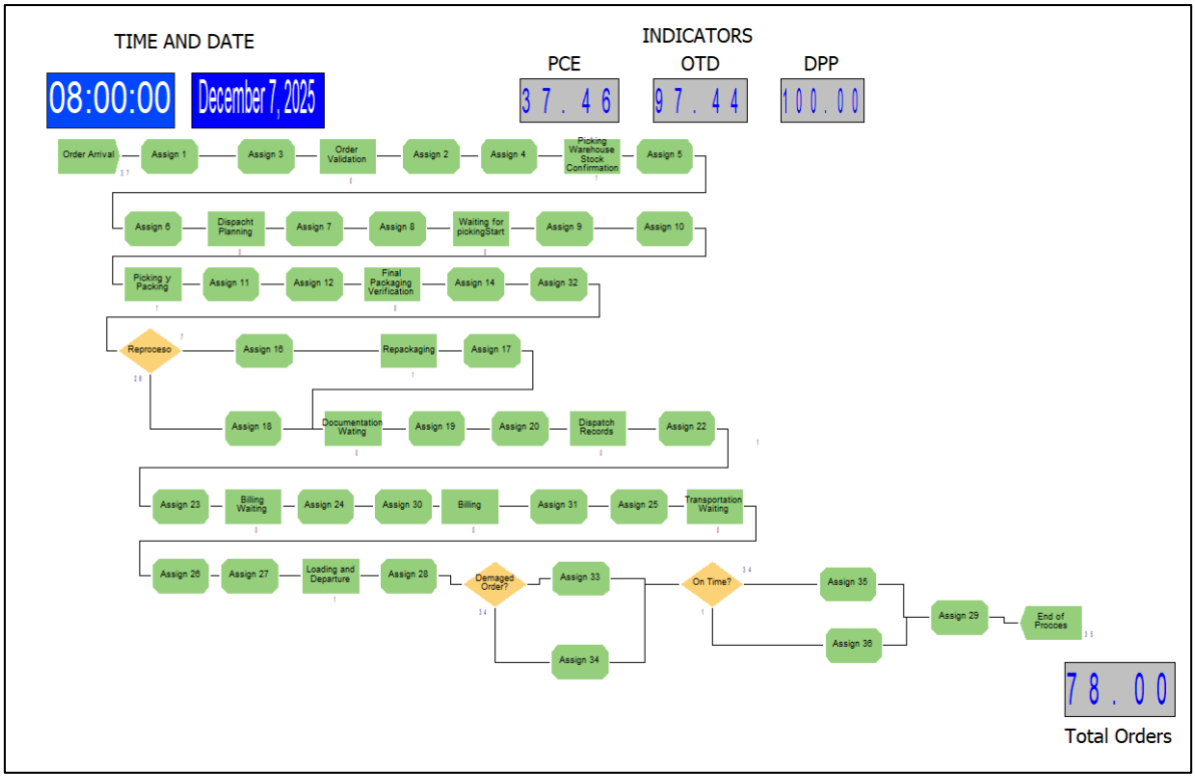


Figure 6: Simulation model in Arena.

Validation Logic & Figure 6: **Figure 6** illustrates the hybrid validation protocol, detailing how empirical data from the physical pilot calibrates the digital twin. Unlike theoretical models, this simulation interacts directly with the pilot results: real-world processing times (picking and packing) were collected and subjected to Goodness-of-Fit tests using the Input Analyzer. The data followed specific statistical distributions (e.g., Beta for picking tasks, Error < 0.005), ensuring that the digital model faithfully replicates the variability of the physical operation.

Following these calibrated simulation runs shown in **Figure 6**, the resulting performance metrics for both the As-Is and To-Be scenarios are consolidated in **Table 2**. This table provides a clear comparison of the simulated indicators, quantifying the improvements in OTD, DPP, PCE, and the overall Perfect Order Rate achieved through the proposed Lean–Digital model.

Table 2: Comparative table of results

Indicator	Unit	As-Is	To-Be (Simulated)	Variation (%)
On-Time Delivery (OTD)	%	91.35%	94.15%	3.07%
Defect-Free Products (DPP)	%	95.12%	98.89%	3.96%
Process Cycle Efficiency (PCE)	%	5.00%	37.00%	640%
Perfect Order Rate (POR)	%	82.51%	88.30%	7.02%

Spatial Optimization & Slotting (Flex Sim)

To isolate the impact of physical layout constraints, a comparative spatial analysis was conducted using Flex Sim. First, an "As-Is" model was built using the original baseline data, identifying that excessive travel distances and aisle congestion were the structural causes of the low initial efficiency (5% PCE).

Subsequently, the "To-Be" model was developed by integrating the optimized operational parameters already validated in the Arena simulation (such as reduced processing times and error rates) to specifically evaluate the impact of spatial interventions. This improved model validated structural corrections: the removal of an obstructive rack to widen the main aisle and the unification of the reception area to establish a continuous flow. Additionally, a standardized route was assigned to the operators based on the ABC-based Slotting Strategy. The comparison between both Flex Sim models demonstrated that these spatial reconfigurations reduced the Total Waiting Time from 931 minutes to just 75 minutes per cycle. As shown in Table 3, this drastic reduction in non-value-added travel was the primary driver for the Process Cycle Efficiency (PCE) surge from 5% to 42% in the spatially optimized scenario. The spatial effects of these layout modifications were evaluated using Flex Sim, and the resulting As-Is and To-Be configurations are depicted in **Figure 7**, which illustrates the impact of the slotting and aisle-reconfiguration interventions on the operational flow.

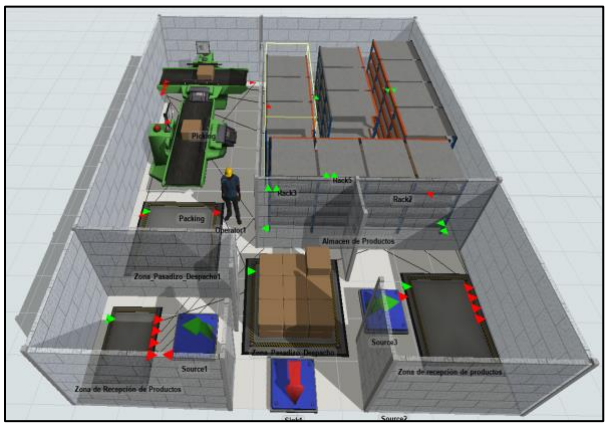
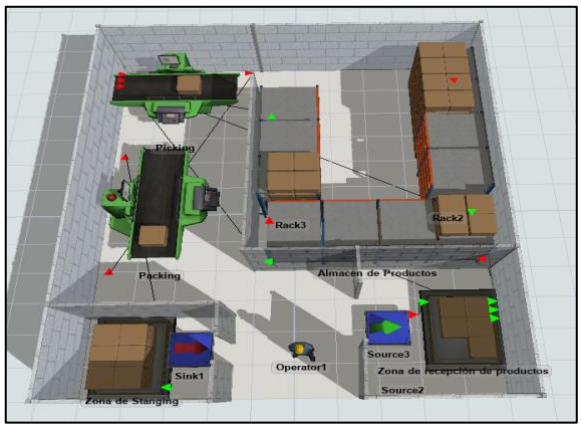
Initial situation (As-Is)	Status before Slotting (To-Be)
	

Figure 7: Flex Sim simulation models: Comparative spatial layouts ("As-Is" vs. "To-Be") validating the foundation for the ABC Classification Slotting Strategy. (https://goo.su/e2zUM)

Following the spatial reconfigurations illustrated in **Figure 7**, **Table 3** presents the quantitative comparison between the As-Is and the optimized To-Be layout. The table consolidates the Flex Sim results, highlighting the substantial reductions in Process Cycle Time and Total Waiting Time, as well as the resulting improvement in Process Cycle Efficiency achieved through the ABC-based slotting and layout adjustments.

Table 3: Comparative table of results of the pilot test

Indicator	Unit	As-Is	Results	Variation (%)
Process Cycle Time (PCT)	minutes	985.38	129.20	87%
Total Waiting Time (TWT)	minutes	931.38	75.20	92%
Process Cycle Efficiency (PCE)	%	5%	42%	740%

5. Results and Discussion

The results of this study demonstrate that the integration of low-cost Lean practices, including 5S, ABC slotting, cycle counting, and poka-yoke mechanisms, generated significant improvements in warehouse performance. This is reflected in the increase of On-Time Delivery from 91.35% to 94.15%, the rise of Defect-Free Products from 95.12% to 98.89%, and the substantial growth of Process Cycle Efficiency from

5% to 37% under simulation. These outcomes align with prior evidence on the capacity of these techniques to reduce search times, enhance throughput, and minimize operational errors [14], [15], [17], [19].

When benchmarked against international logistics quality standards, these findings can be interpreted through the lens of the Perfect Order Rate (POR), a composite indicator that requires simultaneous compliance in timeliness, product integrity, documentation accuracy, and quantity correctness [4]. For this calculation, the simulated post-improvement values for On-Time Delivery (94.15%) and Defect-Free Products (98.89%) were integrated with the company's historical data regarding Order Completeness (96.28%) and Documentation Accuracy (98.44%). Under this framework, the projected POR of approximately 88.3% for the improved scenario exceeds typical performance levels reported for small and medium health-sector enterprises in Latin America, yet remains below the optimal thresholds suggested in the global literature. This gap reflects the structural complexity of medical supply chains, where stringent handling requirements, high traceability demands, and intensive documentation processes interact with persistent variability in upstream regulatory and logistics conditions [6], [7], [8].

This dynamic is further shaped by the country's strong dependence on imported medical inputs, widely documented as a factor that exposes firms to supplier concentration risks, irregular import cycles, and externally imposed regulatory constraints that directly affect operational reliability [16], [9]. The economic relevance of the problem becomes evident in the estimated losses for the analyzed period (S/ 159,382.34), equivalent to approximately 6.54% of total sales. This is consistent with studies showing that extensive product portfolios, intermediary-heavy distribution structures, and limited standardization amplify operational costs and logistics errors [18], [12].

Moreover, recent research on staged digitalization strategies in warehouse environments indicates that the most impactful improvements for small and medium enterprises arise from tactical interventions—such as visual control systems, integrated dashboards, and lightweight technologies—rather than from full-scale automation, reinforcing the rationale behind the present intervention [17], [13]. Taken together, the findings confirm that a coordinated adoption of standardization practices, visual management, strategic inventory allocation, and preventive error-proofing can simultaneously enhance efficiency, accuracy, and overall reliability within logistics processes operating under high regulatory demands and substantial dependence on external supply chain factors.

6. Conclusion

The results of this study demonstrate that a data-driven Lean Warehousing model can significantly enhance logistics performance in a medical supply SME. The combined implementation of 5S, ABC-based slotting, cycle counting, and poka-yoke mechanisms led to measurable operational gains, including improvements in On-Time Delivery from 91.35% to 94.15%, Defect-Free Products from 95.12% to 98.89%, and Process Cycle Efficiency from 5% to 37%. These advances increased the projected Perfect Order Rate from 82.51% to 88.30%, confirming that structured warehouse organization, disciplined standardization, and digital error-proofing are effective enablers of reliability in regulated supply chains. The study highlights the importance of integrating Lean tools into a coherent framework rather than deploying them independently, as their combined effect strengthens visibility, inventory accuracy, and workflow stability. Academically, this research contributes a validated Lean–Digital model tailored to SMEs in the medical supply sector, addressing a gap in the literature on warehouse optimization in emerging markets. Future research should explore long-term adoption through longitudinal studies, evaluate transferability to pharmaceutical and laboratory supply chains, and incorporate real-time digital enablers to support more resilient and data-driven warehouse operations.

Appendix 1: Comparative table of the components of the Proposal Model vs. Literature Review

	Perfect Order and Its Components: Application for Deliveries of Fast Moving Consumer Goods to Retail Stores.		Empirical Validation of a Lean Warehouse Model Using SLP, 5S and ABC in a Peruvian Distribution SME.		A Study on the Application of Logistics Automation in the Healthcare Industry: Exploratory Qualitative Research.		Using Lean Six Sigma to Redesign the Supply Chain to the Operating Room Department of a Private Hospital to Reduce Associated Costs and Release Nursing Time to Care..		Poka Yoke in Smart Production Systems with Pick-to-Light Implementation to Increase Efficiency.		Slotting Optimization Model for a Warehouse with Divisible First-Level Accommodation Locations		**Contribution of this research**	
	Paták, M., Branská, L. & Pecinova, Z. (2020).		Vidal-Palacin, J. & Barrera-Córdova, J. (2025)		Kwak, H., Kim, T.-Y. & Lee, D.-H. (2025).		O'Mahony, L., McCarthy, K., O'Donoghue, J., Teeling, S. P., Ward, M. & McNamara, M. (2021).		Trojanowska, J., Husár, J., Hrehova, S. & Knapčíková, L. (2023).		Viveros, P., González, K., Mena, R., Kristjanpoller, F. & Robledo, J. (2021).		Daza-Llanos, C., Rojas-Lavado, S., Ruiz-Ruiz, M.F. & Quiroz-Flores, J.C.	
Component 0: Problem diagnosis – Literature review – Engineering tools: ABC–PQ, VSM – Case study: Order fulfillment process in a distribution center – Metrics: POR, OTD, DPP, PCE	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓
	Engineering Tool		Engineering Tool	✓	Engineering Tool		Engineering Tool	✓	Engineering Tool		Engineering Tool		Engineering Tool	✓
	Case Study		Case Study	✓	Case Study		Case Study	✓	Case Study		Case Study	✓	Case Study	✓
	Metric	✓	Metric	✓	Metric		Metric		Metric		Metric		Metric	✓
Component 1: Operational synchronization – Engineering tools: 5S, ABC Cycle Counting, Slotting – Case study: Warehouse flow optimization – Metrics: POR, OTD, PCE	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓
	Engineering Tool		Engineering Tool	✓	Engineering Tool		Engineering Tool	✓	Engineering Tool		Engineering Tool	✓	Engineering Tool	✓
	Case Study		Case Study	✓	Case Study	✓	Case Study	✓	Case Study		Case Study	✓	Case Study	✓
	Metric	✓	Metric	✓	Metric		Metric		Metric		Metric		Metric	✓
Component 2: Zero-defect fulfilment – Engineering tools: Poka-Yoke, Double Scanning, Physical Templates, 5S – Case study: Packing quality assurance – Metrics: DPP, Claims, Reprocessing	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓	Literary Review	✓
	Engineering Tool		Engineering Tool		Engineering Tool		Engineering Tool		Engineering Tool	✓	Engineering Tool		Engineering Tool	✓
	Case Study		Case Study		Case Study		Case Study		Case Study	✓	Case Study		Case Study	✓
	Metric	✓	Metric	✓	Metric		Metric		Metric	✓	Metric		Metric	✓
	Total	6	Total	10	Total	4	Total	7	Total	6	Total	6	Total	12

References

[1] INEI, "Informe Técnico: Evolución de las Exportaciones e Importaciones – Mayo 2025," May 2025. Available: https://www.gob.pe/institucion/inei/noticias/1206511-exportaciones-crecieron-7-7-en-mayo-y-acumulan-avance-de-16-2-en-el-2025

[2] Organización Panamericana de la Salud (OPS), "Informe anual 2024: Avanzar en la salud para promover la seguridad regional," May 2025. Available: https://www.paho.org/es/noticias/15-5-2025-informe-anual-2024-ops-destaca-logros-seguridad-sanitaria-americas (accessed December 1, 2025)

[3] Sociedad de Comercio Exterior del Perú (ComexPerú), "Reporte de desempeño logístico: Infraestructura insuficiente y cuellos de botella," 2024. Available: https://logistica360.pe/balance-del-sector-logistico-2024-innovacion-y-desafios-en-peru/ (accessed November 1, 2024)

[4] M. Paták, L. Branská, and Z. Pecinová, "Perfect Order and Its Components: Application for Deliveries of Fast Moving Consumer Goods to Retail Stores," Apr 2020, doi: 10.5755/j01.ee.31.2.22480

[5] J. Vidal-Palacin and J. Barrera-Córdova, "Empirical Validation of a Lean Warehouse Model Using SLP, 5S and ABC in a Peruvian Distribution SME," 2025, https://doi.org/10.14445/23499362/IJIE-V12I1P106

[6] J. E. Tcheng et al., "The Medical Device Unique Device Identifier as the Single Source of Truth in Healthcare Enterprises – Roadmap for Implementation of the Clinically Integrated Supply Chain," Dec 2021, doi: 10.2147/MDER.S344132

[7] C. Amaral, M. Paiva, A. R. Rodrigues, F. Veiga, and V. Bell, "Global Regulatory Challenges for Medical Devices: Impact on Innovation and Market Access," Oct 2024, doi: 10.3390/app14209304

[8] L. O'Mahony et al., "Using Lean Six Sigma to Redesign the Supply Chain to the Operating Room Department of a Private Hospital to Reduce Associated Costs and Release Nursing Time to Care," Oct 2021, doi: 10.3390/ijerph182111011

[9] M. A. Alrasheed, A. U. Rehman, and I. M. Alharkan, "Developing an Efficient Model for Online Grocery Order Fulfillment," May 2024, doi: 10.3390/su16104270

[10] D. D. Rajpuriya, "Regulatory Harmonization: Streamlining Global Pharmaceutical Validation Practices," Jan 2025, doi: 10.32628/CSEIT251112118

[11] H. Kwak, T.-Y. Kim, and D.-H. Lee, "A Study on the Application of Logistics Automation in the Healthcare Industry: Exploratory Qualitative Research," Aug 2025, doi: 10.3390/eng6090205

[12] V. Lukinskiy, V. Lukinskiy, D. Ivanov, B. Sokolov, and D. Bazhina, "A probabilistic approach to information management of order fulfilment reliability with the help of perfect-order analytics," Feb 2023, doi: 10.1016/j.ijinfomgt.2022.102567

[13] B. Unhelkar et al., "Enhancing supply chain performance using RFID technology and decision support systems in the industry 4.0—A systematic literature review," Nov 2022, doi: 10.1016/j.jjimei.2022.100084

[14] B. Kanabar, K. G. Piparva, D. Pandya, and R. B. Kanabar, "The Impact and Challenges of the Implementation of 5S Methodology in Healthcare Settings: A Systematic Review," Jul 2024, doi: 10.7759/cureus.64634

[15] P. Viveros, K. González, R. Mena, F. Kristjanpoller, and J. Robledo, "Slotting Optimization Model for a Warehouse with Divisible First-Level Accommodation Locations," Jan 2021, doi: 10.3390/app11030936

[16] I. R. Destro, F. H. Staudt, K. Somensi, and C. Taboada, "The impacts of inventory record inaccuracy and cycle counting on distribution center performance," 2023, doi: 10.1590/0103-6513.20220077

[17] J. Trojanowska, J. Husár, S. Hrehova, and L. Knapčíková, "Poka Yoke in Smart Production Systems with Pick-to-Light Implementation to Increase Efficiency," Oct 2023, doi: 10.3390/app132111715

[18] J. Bolaños-Zuñiga, M. A. Salazar-Aguilar, and J. A. Saucedo-Martínez, "Solving Location Assignment and Order Picker-Routing Problems in Warehouse Management," Jul 2023, doi: 10.3390/axioms12070711

[19] A. Adeodu, R. Maladzhi, M. G. Kana-Kana Katumba, and I. Daniyan, "Development of an improvement framework for warehouse processes using lean six sigma (DMAIC) approach. A case of third party logistics (3PL) services," Apr 2023, doi: 10.1016/j.heliyon.2023.e14915.

Biographies

Carlos Daza-Llanos (*2002) Carlos Daza-Llanos is an Industrial Engineering student at the University of Lima, ranking in the top tenth of his cohort. He has professional experience in corporate finance and holds certifications in Supply Chain Management and Project Management. His interests focus on continuous improvement, operational analysis, and logistics optimization.

Said Rojas-Lavado (*2002) Said Rojas-Lavado is an Industrial Engineering student at the University of Lima, positioned in the top third of the Faculty of Engineering. He has academic and professional experience in supply chain planning, procurement, and logistics operations. His areas of interest include purchasing management, demand planning, and distribution logistics across industrial sectors.

Marcos Fernando Ruiz-Ruiz holds a PhD in Education Sciences from the Pontifical Catholic University of Peru. He is an Industrial Engineer from the University of Lima and holds a master's degree in education from the University of Piura. He is a Renacyt-accredited researcher and completed an Expert Diploma in Future Studies at UNED (Spain). His research focuses on structural analysis, foresight, and gender approaches within STEM education.

Juan Carlos Quiroz-Flores holds a PhD in Business Administration from the Universidad Nacional Mayor de San Marcos. He is an Industrial Engineer from the University of Lima, an MBA graduate from ESAN University, and a Black Belt in Lean Six Sigma from the Peruvian University of Applied Sciences. He is a professor and research advisor at the University of Lima. His professional background spans more than 25 years in supply chain management, lean operations, process improvement, and productivity enhancement.

8th Conference on Production Systems and Logistics

Challenges And Opportunities In Remanufacturing: A Comprehensive Meta-Analysis

Jan Koller[1], Jonas Weidner[2], Vihasith Ganga[1,3], Frank Döpper[1,2]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Bayreuth, Germany*
[2]*University of Bayreuth, Chair of Manufacturing and Remanufacturing Technology, Bayreuth, Germany*
[3]*Hof University, Hof, Germany*

Abstract

With global waste volumes projected to rise to 3.4 billion tons by 2050 and resource scarcity intensifying, the demand for circular products and sustainable business models is growing. In this context, remanufacturing is a key element of the circular economy, offering a promising end-of-life strategy that enables a new life cycle for products and components. Remanufacturing reduces natural resource consumption and minimizes waste by restoring used products to like-new or better condition, providing significant environmental and economic benefits. Despite these advantages, the full potential of remanufacturing remains unrealized due to various challenges. This paper presents a qualitative meta-analysis of 46 publications to investigate the primary challenges and opportunities associated with remanufacturing. Key challenges and opportunities were systematically analyzed and categorized using a structured coding approach supported by MAXQDA software. The findings reveal various challenges and opportunities, underscoring the complexity of remanufacturing across different industries, and contribute indirectly to address existing research gaps. Industry 4.0 technologies for their realization are described to exploit the potential of remanufacturing opportunities.

Keywords

Remanufacturing; Meta-Analysis; Circular Economy; Sustainability; Challenges and Opportunities

1. Introduction

The prevailing linear economy in industrial production relies on the continuous extraction of raw materials, their transformation into products, and their eventual disposal as waste or emissions, often described as a "take-make-dispose" pattern [1]. This results in significant consequences, including interconnected and cascading effects of climate change, environmental pollution, and biodiversity loss, collectively leading to a "triple planetary crisis" [2]. In response to the need for a transition to a circular economy, remanufacturing has gained growing interest from industry, governments, and science [3,4]. To promote broader industry adoption and optimize remanufacturing processes it is essential to identify and analyze the key challenges and opportunities in remanufacturing.

This paper seeks to consolidate findings from existing research, bringing them to a uniform level of abstraction. By organizing key insights into thematic clusters, it offers a comprehensive overview of the current challenges and opportunities in remanufacturing. These findings will also highlight emerging trends and outline future research directions.

DOI: https://doi.org/10.15488/21026
ISSN: 2701-6277

2. State of the Art

2.1 Remanufacturing

Remanufacturing is an industrial process that restores used products to a quality level equal to or better than new products, accompanied with a full warranty [5–7]. Remanufacturing is a key element of the circular economy, offering ecological, economic, and social benefits compared to other value retention processes such as repair or recycling [6]. Starter motors from the automotive industry serve as an example of how remanufacturing can drastically reduce resource consumption, achieving up to 90% savings in energy and approximately 55% in material usage compared to the production of new starter motors [7,8]. The result is cost savings, with remanufactured products typically priced up to 40% lower than new ones [9].

Figure 1 shows the extended remanufacturing process chain. The center consists of the main process steps of remanufacturing according to STEINHILPER: (1) disassembly, (2) cleaning, (3) inspection and sorting, (4) reconditioning and/or replacing parts, and (5) reassembly [7]. For more complex products, such as electronics, an entrance diagnosis may be required to evaluate the condition of the incoming products [10]. Quality checks are conducted throughout the main process steps to ensure reliability.

Upstream, the process chain starts with the procurement of used parts, which are pre-sorted and stored temporarily. Parts or components that are sorted out during the remanufacturing process are sent for recycling. Therefore, some new parts and/or spare parts must be supplied for reassembly, which are procured and stored in parallel. Once the remanufactured products pass the final testing, they are packaged and prepared for shipment as part of the downstream process. [7,11]

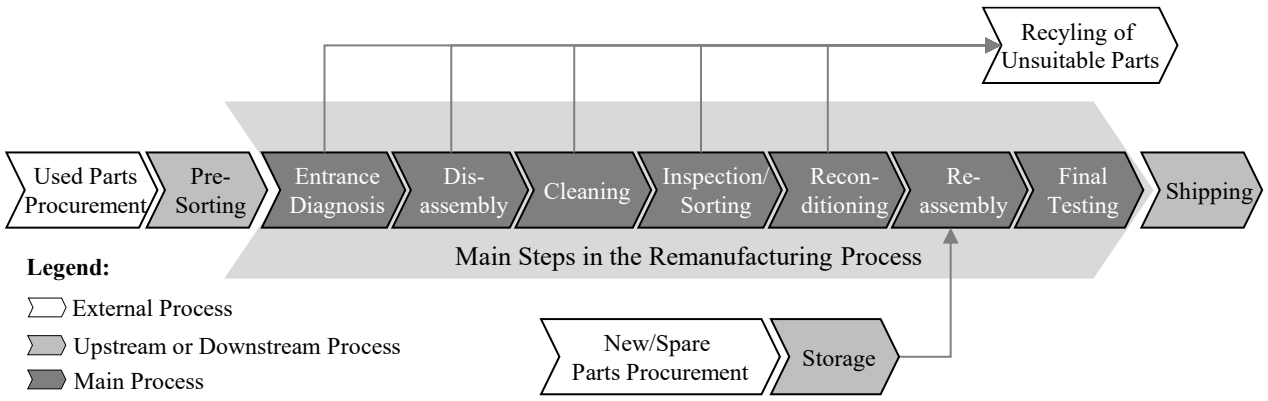


Figure 1: Extended remanufacturing process chain, based on [7,11]

Industries such as automotive, aerospace, electronics, and heavy machinery have established remanufacturing practices, with components like engines and transmissions commonly remanufactured [11]. Through remanufacturing, companies provide consumers with high-quality, cost-effective, and sustainable products. However, the full potential of remanufacturing remains unrealized despite its benefits [12].

2.2 Meta-Analysis

A meta-study systematically evaluates and synthesizes multiple existing studies, often referred to as a secondary analysis or systematic review. Its primary goal is to create coherence from a heterogeneous body of research, providing a unified context. Meta-analysis, a specific form of meta-study, quantitatively combines the results of various studies to identify overall effects and statistical significance. However, it is limited to quantitative data and may lead to publication bias due to selective reporting. [13,14]

Both meta-studies and meta-analyses follow a structured process. COOPER & HEDGES outline six key stages: problem formulation, comprehensive literature search, data evaluation and coding, data analysis and interpretation, and finally, presentation of results [15]. In this process, ensuring a broad literature search and reliable data extraction is critical for drawing meaningful conclusions.

Various methods are used to conduct these studies. Narrative reviews, though flexible, can be biased as they are often unsystematic and may overlook key studies [16]. Meta-synthesis is suited for qualitative data, focusing on interpreting and understanding complex phenomena [17]. Qualitative content analysis organizes text data by identifying, coding, and categorizing relevant passages, allowing for a structured examination of content [14].

Given its systematic approach and ability to integrate diverse studies, this paper employs qualitative meta-analysis, which bridges qualitative and quantitative methods, to provide a comprehensive assessment of the challenges and opportunities in remanufacturing.

3. Qualitative Meta-Analysis on the Challenges and Opportunities in Remanufacturing

3.1 Data Collection through Literature Review

A systematic literature review forms the foundation for the subsequent qualitative meta-analysis by identifying relevant publications. The literature review follows a structured procedure, based on the methodology developed by BRENDEL ET AL. and adapted by KOLLER ET AL., consisting of the planning, conducting, and reporting phases [18,19]. The guiding research question is: "*What are the challenges and opportunities in remanufacturing?*"

The search was conducted using the electronic databases *Web of Science* and *Google Scholar*. *Web of Science* provided access to high-quality, peer-reviewed articles, while *Google Scholar* expanded the search to include gray literature such as government reports. The search was restricted to titles, as full-text searches often generate an excessive number of irrelevant results [18]. Different spellings of the keywords "Challenges", "Issues", "Barriers", "Opportunities", "Trends", and "Potentials", as well as various spellings of "Remanufacturing", were combined in a search query using boolean operators.

The literature review process is documented using the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement" [10]. As shown in Figure 2, the flow diagram outlines the stages of the review, from identification and screening to eligibility and inclusion. The initial search query identified 174 records published since 2011. After removing 68 duplicates, 106 records remained for further screening. During this phase, six studies were excluded based on formal criteria, such as language or accessibility, leaving 100 studies for the eligibility stage. After a detailed assessment, 54 studies were excluded — 27 based on their titles, 23 on abstracts, and four after full-text review. Finally, 46 studies were included in the qualitative synthesis.

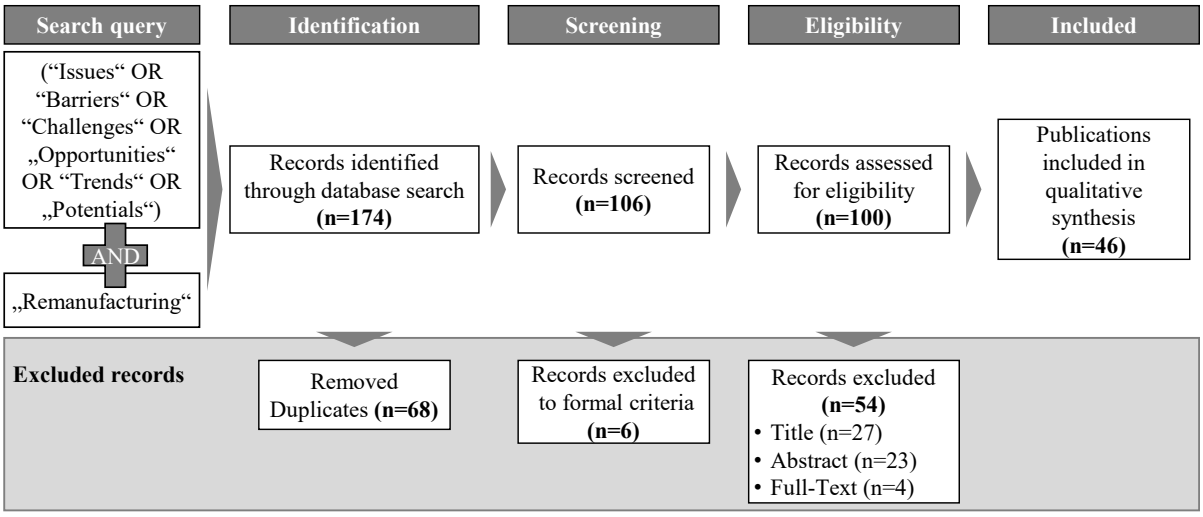


Figure 2: PRISMA flow diagram of the literature review, based on [20]

3.2 Qualitative Meta-Analysis

In the next step, challenges and opportunities in remanufacturing are evaluated based on the 46 publications identified through the literature review. The analysis process follows the methodology of KUCKARTZ [21]. Starting with the initial text work, relevant passages from the selected studies were highlighted. Deductively, two main categories were established based on the research question: “Challenges” and “Opportunities”. Next, relevant text passages were coded with these main categories, which allowed for the compilation of these passages according to their thematic content. The statements varied in length, ranging from single words to entire paragraphs, all of which were systematically categorized. Subsequently, subcategories were inductively developed from the material, facilitating a more detailed analysis. Both specific and abstract challenges were identified and integrated into a hierarchical categorization system. This approach enabled the standardization of the information and the reduction of the text to its essential aspects. Figure 3 illustrates the procedure of the qualitative content analysis.

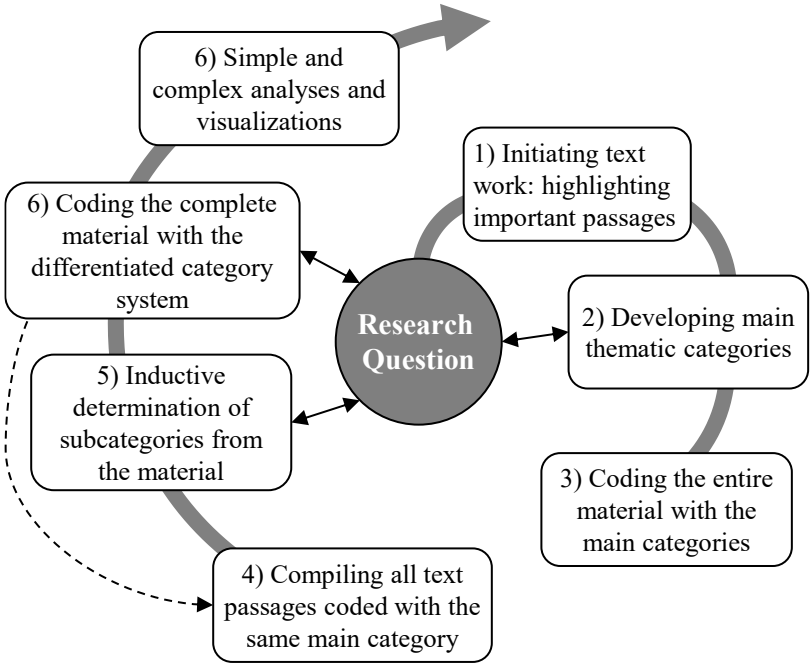


Figure 3: Qualitative content analysis methodology, based on [21]

To manage the volume of data and ensure consistency throughout the coding process, the analysis was supported by the software MAXQDA Analytics Pro 2020. The software was used for both coding text fragments and organizing the categories. Over 1,600 relevant text segments were identified across the 46 studies, which were then systematically coded into the two main categories and their respective subcategories. The software’s functionalities, such as the Creative Coding tool, facilitated the synthesis of related codes and helped construct a coherent, hierarchical system of categories. This process ensured that similar subcategories, like “Product complexity” and “Variety of components”, were merged into broader themes like “Product design”, thereby streamlining the analysis.

4. Results

The results are primarily shaped by the developed category system and the individual statements coded within it. Using the developed categories, the content was structured to create a framework for both the challenges and opportunities in remanufacturing along with a numerical representation of their frequency in the literature. This allows for organizing the relevant information from all studies at a consistent level of abstraction, making it easier to analyze the data comprehensively. Due to space constraints and the large number of studies examined, only a few representative references are cited in this paper to illustrate the

findings in each category.[1] This ensures clarity in presenting the findings while maintaining a focus on key examples to support the overall results.

4.1 Challenges in Remanufacturing

In the process of refining the category system to structure the various challenges in remanufacturing, four main categories were established. These categories distinguish between the challenges encountered during three key phases of remanufacturing: acquisition of parts, the remanufacturing process itself, and the distribution of remanufactured products. Furthermore, higher-level aspects were identified.

During the initial coding phase, a total of 235 categories were developed through open coding by reducing the individual statements from the studies. Since this initial level of categorization had a relatively low level of abstraction, the categories were subsequently refined in a multistep process, resulting in the identification of 23 subcategories that describe the specific challenges encountered in remanufacturing. These subcategories were then assigned to one of the four main categories. Figure 4 illustrates the breakdown of the challenges according to the main categories, with the percentage values representing the proportion of studies that addressed each challenge.

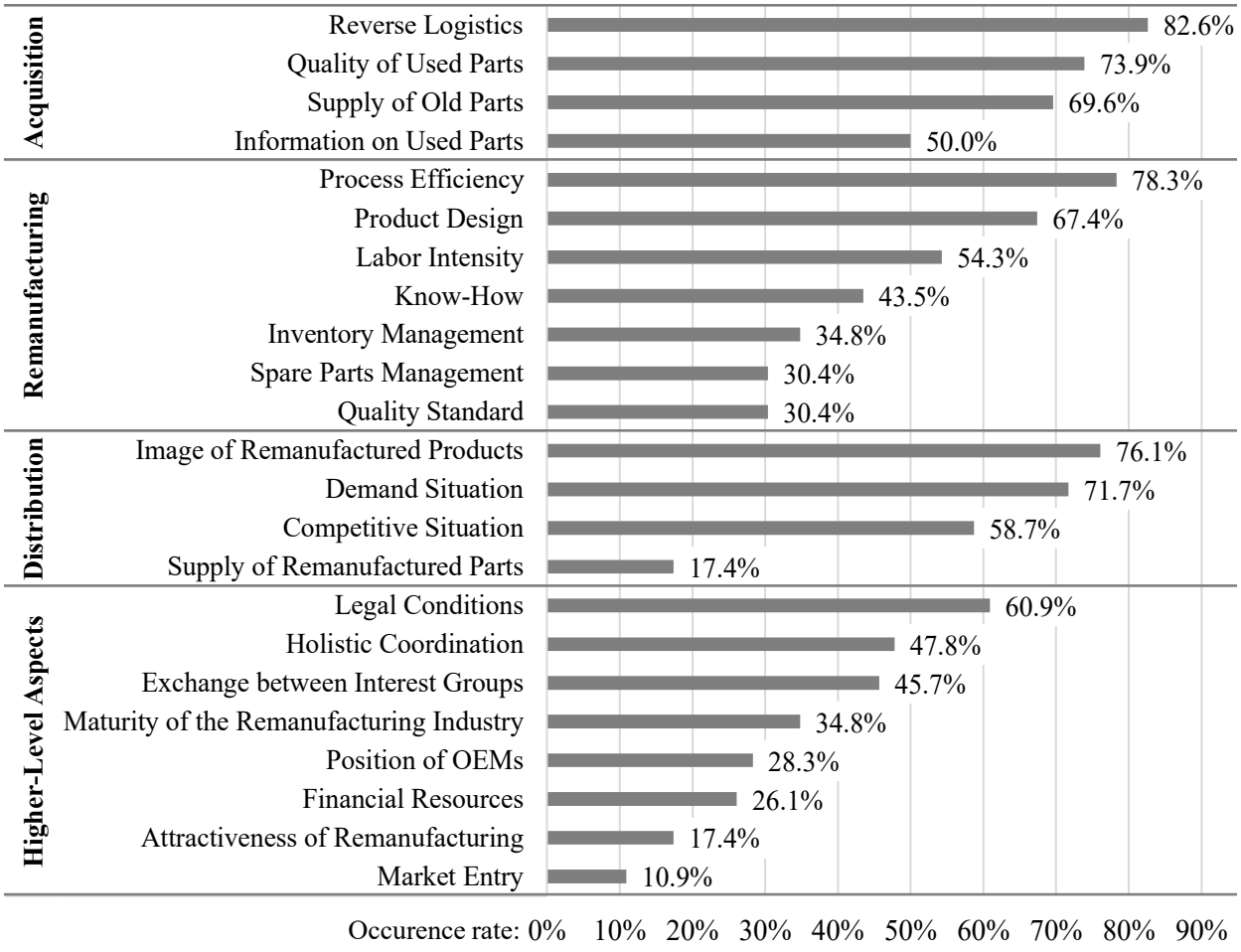


Figure 4: Challenges in remanufacturing and their occurrence rate in the literature (n=46)

The analysis revealed that most challenges align with recurring themes, with only a few outliers. High percentages show that most text segments focus on a consistent set of issues, reinforcing the reliability of the categorization and abstraction, with few instances below 30%. Among the 23 challenges, the top five, which are mentioned in over 70% of the literature, are listed below.

[1] A complete list of all analyzed studies with references to the source material for each category can be requested from the authors.

- **Reverse Logistics**: A major challenge is the unpredictability of quantity, timing, and condition of returned products [22,23], which makes it difficult to plan inventory and production schedules. Therefore, organizing efficient collection and transporting used products from diverse locations [24] can be complex and costly [25], particularly if returns are low in volume or dispersed.
- **Process Efficiency**: Lack of specific technologies for remanufacturing [26] and high process costs due to the unpredictability of processing and waiting times [27] and variable lead times lead to reduced process efficiency.
- **Image of Remanufactured Products**: In many cases, there is a low level of acceptance of the remanufactured products [28], which is due to customers' lack of knowledge about the expected quality and price [29]. Additionally, the presence of unreliable remanufactured products contributes to a negative reputation [30].
- **Quality of Used Parts**: Remanufacturers encounter challenges in obtaining used products in good condition [4] and face significant unpredictability in sourcing used parts due to the wide variation in the usage and condition of returned products [31].
- **Demand Situation**: Low purchasing interest [24], lack of a specific organized market [29], and absence of distribution and marketing channels [32] make purchasing remanufactured products partially impossible [25].

The specific challenges vary depending on a company's position in the value chain. For instance, product returns can be especially challenging for companies whose products are embedded within other products, such as electric motors within electric vehicles. In parallel to the existing challenges, there are also opportunities for remanufacturing, which are examined in more detail below.

4.2 Opportunities in Remanufacturing

The identified opportunities in remanufacturing are future-oriented and thus not guaranteed to materialize or produce positive effects. The goal was to group these statements into congruent subcategories to reduce the diversity of insights to their essential elements. During the coding process and subsequent category development, 201 initial categories were identified to describe the individual statements. Due to their high level of specificity, these were consolidated into 18 subcategories, forming the final category system. The subcategories were kept at the same level to ensure equal emphasis on each opportunity and avoid an implied ranking or prioritization. Figure 5 illustrates the overview of opportunities in remanufacturing and their occurrence rate in the literature.

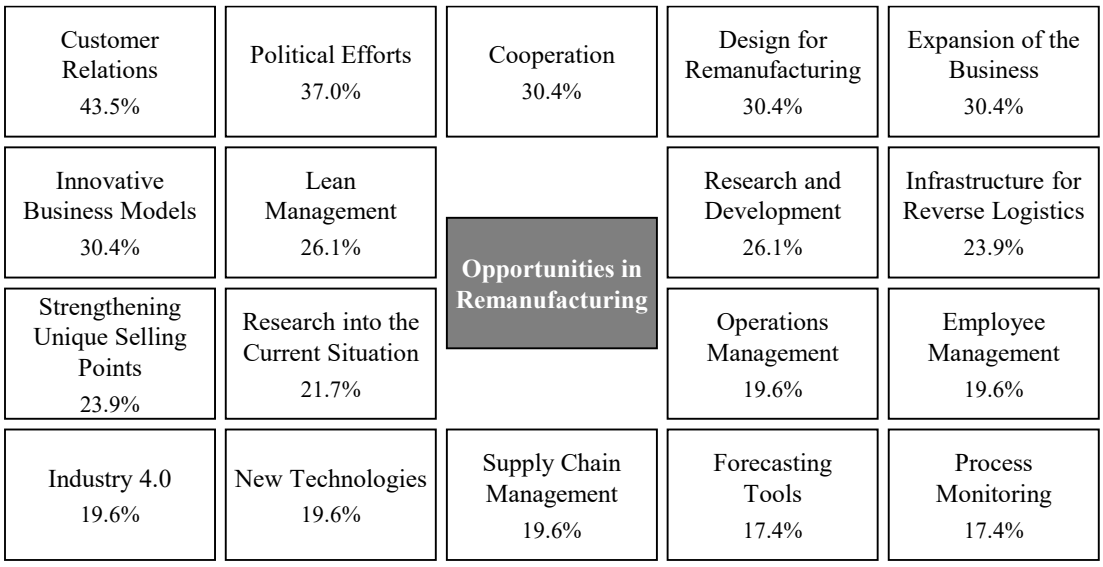


Figure 5: Opportunities in remanufacturing and their occurrence rate in the literature

The six subcategories, mentioned in over 30% of the literature, are outlined below.

- **Customer Relations**: By organizing campaigns that promote the purchase or use of remanufactured products [33], sales figures and the accompanying revenue can be increased and additionally consumer awareness should be raised [23] highlighting the economic and ecological benefits [4].
- **Political Efforts**: Government support in the form of incentives or subsidies [34] and collaborative efforts with remanufacturers [35] will have a positive impact. Guidelines for remanufacturing should be developed [36], and outdated regulations should be revised [30].
- **Cooperation**: Strategic partnerships with Original Equipment Manufacturers (OEMs), suppliers, service providers and recycling facilities can optimize resource sharing, and improve access to essential components and technology. Furthermore, remanufacturers can provide user and design feedback to the OEMs. [12]
- **Innovative Business Models**: Business models like Product Service Systems align incentives with remanufacturing and circular practices [37]. These models create a win-win situation by improving the perception of remanufactured products while allowing companies to generate revenues.
- **Design for Remanufacturing**: Designing products for feasible remanufacturing improves disassembly, inspection, and component recovery processes, reducing costs and energy use [38]. Additionally, integrating remanufacturing principles during the design phase not only extends product lifespan but also enables consideration of potential future upgrades [12].
- **Expansion of the Business**: The growing market for remanufactured products [11], driven by the aging of products such as cars [31] or airplanes [4], supports business growth. Technological advancements and the emergence of new fields, e.g. due to the transformation to electromobility, offer a new market potential [31,31].9

The subcategories are based on specific terms from the reviewed studies. However, consolidating these into broader subcategories led to some loss of detail, leaving certain relevant elements only partially addressed. Given the critical importance of digitalizing processes, products, and information exchange across the value chain [40], the following section examines potential applications of digitalization in remanufacturing.

5. Enhancing Remanufacturing Potentials with Digital Technologies

Digitalization could play a central role in overcoming challenges in remanufacturing, extending beyond simply enhancing production efficiency [40]. While the use of Industry 4.0 technologies along the remanufacturing process chain has already been examined [41], the following section will look at the potential in an overarching context based on the categories Smart Data, Smart Interaction, and Smart Operation developed by DILLINGER ET AL. [42].

5.1 Smart Data

Through Smart Data, remanufacturers gain enhanced access to product and usage information, improving insights into return flows and product conditions, which helps in accurate planning and quality assessment.

- **Big Data and Analytics**: By tracking product lifecycle stages [43] and analyzing usage and demand patterns, big data supports remanufacturers in managing uncertainties within scheduling activities [44]. It also offers the capability to attract target consumers [43] and supports condition monitoring.
- **Internet of Things (IoT)**: IoT devices can capture real-time data on product condition and usage to optimize the acquisition strategy [45], and forecast supply and demand for used and remanufactured products. Companies can access used product data from external OEMs [46], determine the actual conditions of products and use this data to evaluate design alternatives for end-of-life products [47].
- **Cloud**: Products and machine related data can be evaluated and stored in real time, enabling data availability independent of location, time and device for an efficient exchange of data [48,49].
- **Cyber Security**: Robust cyber security measures protect sensitive data within the remanufacturing process, safeguarding information integrity [50] and promoting trust among stakeholders and clients.

5.2 Smart Interaction

Smart Interaction enables seamless data exchange and improved transparency between stakeholders in the value chain, supporting coordination for optimal product retrieval timing and remanufacturing readiness.

- **Simulation**: Simulation tools enable remanufacturers to evaluate environmental impacts [51], predict process outcomes, optimize resource use and understand system behaviour under variable conditions. Additionally, digital twins provide a real-time, dynamic representation of physical assets, facilitating ongoing simulation and process adjustments for greater precision and efficiency.
- **Horizontal and Vertical Integration**: Horizontal integration enables faster market responsiveness, improving adaptability of remanufacturers, while vertical integration helps reduce costs, improve quality control, and minimize uncertainties in supply chain management [52].

5.3 Smart Operation

Lastly, Smart Operation leverages digital tools to monitor and adapt remanufacturing processes in real time, facilitating proactive maintenance and resource efficiency, ultimately enabling operational circularity.

- **Additive Manufacturing (AM)**: AM allows for the decentralized on-demand production of spare parts and components, supporting timely repair or reconditioning of worn parts [19,53]. AM is particularly beneficial for complex or custom parts and small series production.
- **Autonomous Processes:** Automation in remanufacturing holds great potential that has not yet been exploited [54]. Robots can be used, for example, for complex or dangerous tasks and to increase efficiency [49]. Additionally, the matrix production approach, enabling flexible product routing, reducing waiting times and enhancing flexibility could be applied in remanufacturing [55].
- **Assistance Systems:** Augmented and Virtual Reality can be used to train the workforce [56] by providing real-time information about the work environment [49] and guiding them on disassembly operations [57–59]. Additionally, cobots can be used, e.g. to support disassembly or inspection to increase efficiency and improve occupational safety.

Digitalization enables remanufacturing opportunities by improving operational agility and strengthening the value chain. Furthermore, Industry 4.0 supports innovative business models and scalable expansion into emerging markets like electric motor and battery system remanufacturing in the shift to electromobility [39].

6. Conclusion and Outlook

This paper highlights the current challenges and opportunities of remanufacturing. Using MAXQDA to code the findings from 46 relevant publications identified through a literature review, more than 1,600 relevant text segments were analyzed and key challenges and opportunities in remanufacturing were identified. Furthermore, Industry 4.0 technologies for exploiting the opportunities were presented. While the qualitative meta-analysis offered structured insights, some information loss occurred when harmonizing diverse findings and integrating data. Despite these limitations, the study provides a comprehensive understanding of current remanufacturing challenges and opportunities.

However, deeper investigation into specific challenges and opportunities in remanufacturing remains essential. Future academic work should build on these findings to advance remanufacturing practices, support industry development, and contribute to sustainability. Furthermore, collaboration with industry and policymakers, as well as pilot projects, will be crucial for testing and implementing these approaches.

Acknowledgements

This paper evolved from the research project “REASSERT - Value retention strategies for electric traction motors” and was funded by the by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) with Grant No 19S23004C.

References

[1] Ellen Macarthur Foundation, 2013. Towards the Circular Economy Vol. 1: Economic and business rationale for an accelerated transition.

[2] European Commission, 2022. COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT: Accompanying the document Proposal for a Regulation of the European Parliament and of the Council establishing a framework for setting ecodesign requirements for sustainable products and repealing Directive 2009/125/EC, Brussels.

[3] Fitzsimons, D. European Remanufacturing Council Achievements 2017 – 2019, Brussels.

[4] Parker, D., Riley, K., Robinson, S., Symington, H., Tewson, J., Jansson, K., Ramkumar, S., Peck, D., 2015. Remanufacturing Market Study. European Remanufacturing Network, 145 pp.

[5] British Standards Institution, 2009. BS 8887-2:2009: Design for manufacture, assembly, disassembly and end-of-life processing (MADE). Part 2: Terms and definitions, London.

[6] Deutsches Institut für Normung e. V., 2023. DIN SPEC 91472: Remanufacturing (Reman) - Qualitätsklassifizierung für zirkuläre Prozesse. Beuth Verlag, Berlin.

[7] Steinhilper, R., 1998. Remanufacturing: The ultimate form of recycling. Fraunhofer IRB Verlag, Stuttgart.

[8] Köhler, D.C.F., 2011. Regenerative Supply Chains. Shaker Verlag, Aachen.

[9] Colledani, M., Battaïa, O., 2016. A decision support system to manage the quality of End-of-Life products in disassembly systems. CIRP Annals 65 (1), 41–44.

[10] Freiberger, S., 2007. Prüf- und Diagnostechnologien zur Refabrikation von mechatronischen Systemen aus Fahrzeugen. Shaker Verlag, Düren.

[11] Lange, U., 2017. Resource efficiency through remanufacturing. VDI Zentrum Ressourceneffizienz GmbH (VDI ZRE), Berlin.

[12] Boorsma, N., Peck, D., Fischer, S., Bakker, C., Balkenende, R., 2018. Capabilities required to tackle barriers to remanufacturing, in: Proceedings of the Conference Going Green – CARE INNOVATION 2018, Wien.

[13] Fink, A., 2020. Conducting research literature reviews: From the internet to paper, 5th ed. SAGE, Los Angeles, 280 pp.

[14] Schnepf, J., Groeben, N., 2019. Qualitative Metaanalyse mithilfe computergestützter qualitativer Inhaltsanalyse – am Beispiel von Lokale-Agenda-21-Prozessen. Forum Qualitative Sozialforschung 20 (3).

[15] Cooper, H., Hedges, L.V., 2009. Research synthesis as a scientific process, in: Cooper, H., Hedges, L.V., Valentine, J.C. (Eds.), The Handbook of Research Synthesis and Meta-Analysis, 2 ed. Russell Sage Foundation, New York, pp. 3–16.

[16] Krause, B.J., Khan, C., Antoch, G., 2015. Wie schreibe ich eine gute Übersichtsarbeit? Nuklearmedizin 54 (2), 69–74.

[17] Jensen, L.A., Allen, M.N., 1996. Meta-Synthesis of Qualitative Findings. Qualitative Health Research 6 (4), 553–560.

[18] Brendel, A.B., Trang, S., Marrone, M., Lichtenberg, S., Kolbe, L.M., 2020. What to do for a Literature Review? – A Synthesis of Literature Review Practices, in: AMCIS 2020 Proceedings.

[19] Koller, J., Häfner, R., Döpper, F., 2022. Decentralized Spare Parts Production for the Aftermarket using Additive Manufacturing - A Literature Review. Procedia CIRP 107, 894–901.

[20] Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Medicine 6 (7), 1–6.

[21] Kuckartz, U., 2014. Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung, 2nd ed. Beltz Juventa, Weinheim, Basel.

[22] Casper, R., Sundin, E., 2018. Addressing Today's challenges in automotive remanufacturing. Journal of Remanufacturing 8, 93–102.

[23] Chakraborty, K., Mondal, S., Mukherjee, K., 2019. Critical analysis of enablers and barriers in extension of useful life of automotive products through remanufacturing. Journal of Cleaner Production 227, 1117–1135.

[24] Karvonen, I., Jansson, K., Tonteri, H., Vatanen, S., Uoti, M., 2015. Enhancing remanufacturing – studying networks and sustainability to support Finnish industry. Journal of Remanufacturing 5 (5), 1–16.

[25] Walsh, B., Waugh, R., Symington, H., 2015. Circular Economy Evidence Building Programme: Remanufacturing Study. Oakdene Hollins.

[26] Wei, S., Cheng, D., Sundin, E., Tang, O., 2015. Motives and barriers of the remanufacturing industry in China. Journal of Cleaner Production 94, 340–351.

[27] Kurilova-Palisaitiene, J., Sundin, E., Poksinska, B., 2018. Remanufacturing challenges and possible lean improvements. Journal of Cleaner Production 172, 3225–3236.

[28] Guidat, T., Uoti, M., Tonteri, H., Määttä, T., 2015. A Classification of Remanufacturing Networks in Europe and their Influence on New Entrants. 12th Global Conference on Sustainable Manufacturing. Procedia CIRP 26, 683–688.

[29] Israr, M., Gangele, A., Kripalani, K., 2013. Analysis and Optimization of Remanufacturing in India through Approaches, Potentials & Technical challenges. International Journal of Emerging Technology and Advanced Engineering 3 (9), 362–364.

[30] Zhang, J.-H., Yang, B., Chen, M., 2017. Challenges of the development for automotive parts remanufacturing in China. Journal of Cleaner Production 140, 1087–1094.

[31] Yang, S., 2020. Global Challenges and Market Transformation in Support of Remanufacturing, in: Nasr, N. (Ed.), Remanufacturing in the Circular Economy. John Wiley & Sons, Ltd, Hoboken, pp. 169–210.

[32] United States International Trade Commission, 2012. Remanufactured Goods: An Overview of the U.S. and Global Industries, Markets, and Trade.

[33] Jansson, K., Vatanen, S., Karvonen, I., Behm, K., Waugh, R., Fitzsimons, D., Sundin, E., Parker, D., 2017. D6.3 Targeted Recommendations: For Horizon 2020, grant agreement No 645984. European Remanufacturing Network.

[34] Govindan, K., Madan Shankar, K., Kannan, D., 2016. Application of fuzzy analytic network process for barrier evaluation in automotive parts remanufacturing towards cleaner production – a study in an Indian scenario. Journal of Cleaner Production 114, 199–213.

[35] Shi, J., Zhou, J., Zhu, Q., 2019. Barriers of a closed-loop cartridge remanufacturing supply chain for urban waste recovery governance in China. Journal of Cleaner Production 212, 1544–1553.

[36] Gunasekara, H., Gamage, J., Punchihewa, H., 2020. Remanufacture for sustainability: Barriers and solutions to promote automotive remanufacturing. Procedia Manufacturing 43, 606–613.

[37] Shao, J., Huang, S., Lemus-Aguilar, I., Ünal, E., 2020. Circular business models generation for automobile remanufacturing industry in China: Barriers and opportunities. Journal of Manufacturing Technology Management 31 (3), 542–571.

[38] Lee, C.-M., Woo, W.-S., Roh, Y.-H., 2017. Remanufacturing: Trends and issues. International Journal of Precision Engineering and Manufacturing-Green Technology 4 (1), 113–125.

[39] Roland Berger & CLEPA, 2022. The electrification of light vehicles: Boon or bane for the European aftermarket?

[40] acatech/Circular Economy Initiative Deutschland/SYSTEMIQ, 2020. Circular Business Models: Overcoming Barriers, Unleashing Potentials: Executive Summary and Recommendations, Munich.

[41] Sprenger, K., Klein, J.-F., Wurster, M., Stricker, N., Lanza, G., 2021. Industrie 4.0 im Remanufacturing. I40M 2021 (4), 37–40.

[42] Dillinger, F., Messmer, C., Reinhart, G., 2021. Industrie-4.0-Technologiekreis für produzierende Unternehmen. ZWF 116 (9), 639-343.

[43] Bressanelli, G., Adrodegari, F., Perona, M., Saccani, N., 2018. Exploring How Usage-Focused Business Models Enable Circular Economy through Digital Technologies. Sustainability 10 (3), 639.

[44] Dinis, D., Barbosa-Póvoa, A., Teixeira, Â.P., 2019. Valuing data in aircraft maintenance through big data analytics: A probabilistic approach for capacity planning using Bayesian networks. Computers & Industrial Engineering 128, 920–936.

[45] Fang, C., Liu, X., Pardalos, P.M., Pei, J., 2016. Optimization for a three-stage production system in the Internet of Things: procurement, production and product recovery, and acquisition. Int J Adv Manuf Technol 83 (5-8), 689–710.

[46] Alqahtani, A.Y., Gupta, S.M., Nakashima, K., 2019. Warranty and maintenance analysis of sensor embedded products using internet of things in industry 4.0. International Journal of Production Economics 208, 483–499.

[47] Joshi, A.D., Gupta, S.M., 2019. Evaluation of design alternatives of End-Of-Life products using internet of things. International Journal of Production Economics 208, 281–293.

[48] Bauernhansl, T., Hompel, M. ten, Vogel-Heuser, B., 2014. Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung · Technologien · Migration. Springer Vieweg, Wiesbaden.

[49] Rüßmann,, M., Lorenz, M., Gerbert, P.M., Waldner, M., Justus, J., Engel, P., Harnisch, M., 2015. Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries. Boston Consulting Group, 14 pp.

[50] Rao, U., 2023. Overview of Cyber Security. IJARSCT, 47–51.

[51] Liu, M., Chen, L., Sheng, X., Yang, Y., Yu, F., Li, Y., Yuan, X., Li, Y., Wang, Q., Ma, Q., 2023. Dynamic simulation of life cycle environmental benefits of remanufacturing asynchronous motors to permanent magnet synchronous motors. Journal of Cleaner Production 426, 138932.

[52] Plum, D.M., Peters, B., Shaw, D., 2017. Balancing the Vertical Integration vs. Out-Sourcing Decision for Quality and Continuity. Strategic Management Quarterly 5 (1&2), 1–4.

[53] Khershed P. Cooper, 2014. Laser-based additive manufacturing: where it has been, where it needs to go, in: Proceedings of SPIE, Vol. 8970.

[54] Nohra, P., Rejeb, H.B., Venkateswaran, S., 2022. Impact of automation during innovative remanufacturing processes in circular economy: a state of the art. http://arxiv.org/pdf/2209.04040v1.

[55] Greschke, P., Schönemann, M., Thiede, S., Herrmann, C., 2014. Matrix Structures for High Volumes and Flexibility in Production Systems. Procedia CIRP 17, 160–165.

[56] Ceruti, A., Marzocca, P., Liverani, A., Bil, C., 2019. Maintenance in aeronautics in an Industry 4.0 context: The role of Augmented Reality and Additive Manufacturing. Journal of Computational Design and Engineering 6 (4), 516–526.

[57] Kerin, M., Pham, D.T., 2020. Smart remanufacturing: a review and research framework. JMTM 31 (6), 1205–1235.

[58] Koller, J., Kemp, D., Dopper, F., 2020. Supporting Disassembly in Remanufacturing with Augmented Reality, in: 2020 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD). 2020 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Marrakech, Morocco. 11/24/2020 - 11/27/2020. IEEE, Piscataway, NJ, pp. 1–8.

[59] S. Butzer, D. Kemp, R. Steinhilper, S. Schötz, 2016. Identification of approaches for remanufacturing 4.0, in: 2016 IEEE European Technology and Engineering Management Summit (E-TEMS). 2016 IEEE European Technology and Engineering Management Summit (E-TEMS), pp. 1–6.

Biography

Jan Koller (*1992) heads the research team “Remanufacturing Operations and Management” at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Bayreuth since 2022. He holds a doctorate from the Faculty of Engineering at the University of Bayreuth and deals with the challenges of remanufacturing in national and international research and industrial projects.

Jonas Weidner (*1997) studied Business Administration and Engineering at the University of Bayreuth and wrote his bachelor's thesis under the supervision of Professor Döpper and Jan Koller at the Chair of Manufacturing and Remanufacturing Technology.

Vihasith Ganga (*1998) is a master’s student in Engineering and Business Administration at Hof University of Applied Sciences, specializing in operational excellence. He holds a bachelor’s degree in mechanical engineering and has gained valuable industry experience in operational excellence. Currently, he is a student assistant at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Bayreuth, and his research areas include production systems, digitalization, and remanufacturing.

Frank Döpper (*1970) is Scientific Director for Manufacturing and Process Technology at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA and holds the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth. For more than twenty years, he worked in technology and investment management and as a site and production manager before he was appointed full professor at the University of Bayreuth.

8th Conference on Production Systems and Logistics

Assessment Of Electric Off-Road Machinery For Industrial Applications In Uruguay

Daniel Jurburg, Marcelo Berglavaz, Justino Romano, Rodrigo Sanchez, Camila Perez, Micaela Montilla

CINOI, School of Engineering, Universidad de Montevideo, Uruguay

Abstract

Uruguay has advanced an agenda centered on energy sovereignty and deep decarbonization, achieving significant progress in electrifying segments of the transport sector such as taxis, light-duty vehicles, buses, and trucks. However, the applicability of electromobility in off-road machinery remains largely unexplored despite its relevance for achieving national climate objectives. In this context, the construction sector and related activities represent a strategic domain for evaluating the potential transition from diesel-powered equipment to battery-electric machinery. This study examines the feasibility, challenges, and opportunities associated with the introduction of battery-electric heavy equipment in these sectors, adopting a multi-method approach. The research began with a multi-stakeholder workshop conducted with machinery users, suppliers, and technical specialists to identify context-specific considerations regarding infrastructure, operational practices, and perceived risks. Building on these insights, a sequence of semi-structured interviews and a targeted survey was carried out to characterize sectoral readiness, operational requirements, and the principal barriers and enablers for technology adoption. Finally, a Total Cost of Ownership (TCO) analysis was developed using a quarry operation as a representative case study to compare battery-electric and conventional diesel equipment under realistic duty cycles and energy cost scenarios. The results reveal that, although infrastructure availability, charging logistics, and initial investment levels remain significant challenges, the preliminary TCO outcomes suggest favorable economic potential for selected equipment categories. These findings provide a foundational evidence base to inform strategic decisions and support prospective initiatives aimed at accelerating the adoption of battery-electric machinery in Uruguay's construction-related industries.

Keywords

Electromobility; Off-road equipment; decarbonization; Battery-electric vehicles; construction sector; energy transition, Latin America, Heavy-duty machinery

1. Introduction

Electrification has become a central pathway for reducing greenhouse gas emissions in the global transport and industrial sectors. Worldwide, the transport sector accounts for roughly one quarter of primary energy consumption and remains heavily dependent on fossil fuels, making its decarbonization both urgent and complex. Over the last decade, Uruguay has undertaken an ambitious energy transition centered on energy sovereignty, diversification of its electricity matrix, and the systematic reduction of carbon-intensive energy sources. This transition has enabled significant advances in the electrification of several on-road transport

DOI: https://doi.org/10.15488/21027
ISSN: 2701-6277

segments, including taxis, light-duty vehicles, public buses, and freight trucks, demonstrating favorable operational and economic outcomes.

Despite this progress, the potential for electrification in heavy-duty off-road machinery—particularly within the construction sector and related industries—remains largely unexplored. These activities represent an important share of energy use and emissions, yet they have received comparatively limited attention in academic research and public policy discussions. The emergence of battery-electric construction machinery, driven by rapid advances in energy storage, powertrain efficiency, and charging technologies, opens new opportunities for reducing operational emissions, improving local air quality, and lowering noise levels in industrial environments. However, its adoption requires overcoming significant challenges related to technology maturity, infrastructure availability, operational constraints, and upfront investment costs.

Given these uncertainties, there is a clear need for context-specific analysis to evaluate the feasibility and implications of introducing battery-electric machinery in Uruguay. This study addresses these gaps by evaluating the feasibility, opportunities, and barriers associated with introducing battery-electric machinery in Uruguay's construction and quarrying sectors. The study draws on a multi-method approach that integrates a literature review, stakeholder engagement, and a Total Cost of Ownership (TCO) assessment, providing the first systematic evidence base to support future decision-making and the potential development of targeted policies in this emerging area.

The remainder of the article is organized as follows. **Chapter 2** provides a concise review of the relevant international and regional literature. **Chapter 3** presents the methodological framework, including stakeholder engagement activities. **Chapter 4** discusses the main insights derived from the workshop, interviews, and survey, focusing on perceived opportunities and barriers. **Chapter 5** examines the results of the Total Cost of Ownership (TCO) analysis using a representative quarry case study. Finally, **Chapter 6** summarizes the key conclusions and outlines future lines of research and development.

2. Literature review

Electrification has emerged as one of the most promising pathways for reducing emissions from the transport sector, which remains particularly difficult to decarbonize due to its structural dependence on fossil fuels and diverse operational requirements [1], [2] & [3]. Across Latin America and the Caribbean, several countries have deployed policies to promote electromobility in areas such as taxi fleets, public buses, urban logistics, and motorcycles. In Uruguay, the significant penetration of renewable energy into the electricity matrix enhances the primary-energy efficiency gains achievable through transport electrification. Nationally and globally, the transport sector accounts for approximately 25% of total energy consumption [4] & [5].

Uruguay's 2005–2030 National Energy Policy—developed through a cross-party agreement—established the foundations for a comprehensive set of demand-side strategies to reduce fossil-fuel dependence [6]. Within this framework, the MOVÉS Program [7], initially financed by the Inter-American Development Bank, introduced targeted incentives for the electrification of specific transport subsectors including taxis, buses, last-mile distribution, and motorcycles. Nonetheless, one major segment remains largely unaddressed: heavy off-road machinery used in construction, forestry, mining, and industrial operations. Although definitions vary across jurisdictions, regional regulatory frameworks typically classify off-road equipment as mobile machinery not intended for passenger or freight transport on public roads, predominantly powered by internal combustion engines [8] & [9]. Importantly, these definitions still explicitly reference combustion technologies, underscoring both the historical scarcity of electric alternatives and the traditional regulatory focus on pollutant-emission control.

The decarbonization of off-road machinery is technologically aligned with the transport sector, and four main propulsion pathways have been identified with distinct levels of maturity and commercial development: battery-electric systems, internal combustion engines powered by green hydrogen, fuel-cell systems using

green hydrogen, and hybrid architectures combining combustion engines with electric drives [10], [11] & [12].

Among these, hydrogen-based technologies are the least commercially advanced due primarily to the absence of a mature supply chain and widespread refueling infrastructure [13]. Furthermore, hybrid machinery presents additional considerations. Local suppliers report that these platforms entail a substantial increase in maintenance complexity due to their combination of mechanical, electrical, and electronic subsystems. This aligns with broader evidence from the vehicle sector, where hybrid and electric systems are associated with longer repair times and more demanding maintenance requirements [14]. For these reasons, the present study focuses exclusively on battery-electric off-road machinery.

In terms of convincing the users to adopt these new technologies, on the one hand, the construction sector faces a unique set of operational and logistical challenges that complicate the adoption of battery-electric machinery. These include highly variable duty cycles, demanding work environments, limited charging infrastructure in remote or temporary worksites, stringent performance requirements, and significant capital costs associated with fleet renewal [15] & [16]. Moreover, the sector is characterized by heterogeneous machinery profiles, fragmented ownership structures, and high sensitivity to project timelines and equipment availability [17]. On the other hand, sectors such as forestry also have a high demand for machine working hours, making it unfeasible to adopt alternatives that require stopping operations for electrical recharging [18] & [19].

Together, this body of literature underscores the relevance—and current gap—in assessing the feasibility of electromobility within heavy-duty off-road applications, particularly in contexts like Uruguay where policy frameworks and renewable electricity supply create favourable conditions for deeper decarbonization. The limited international deployment of electric off-road machinery further highlights the need for country-specific analysis that integrates technological, operational, economic, and regulatory considerations.

3. Methodology

This study employs a multi-method research design aimed at identifying the opportunities, barriers, and feasibility conditions for the adoption of battery-electric heavy machinery in Uruguay's construction and related sectors. The methodological approach triangulates publicly available information, stakeholder perspectives, and empirical economic analysis to develop a comprehensive and context-specific assessment.

The analysis begins by identifying which activity sectors present the greatest potential for electrification and by determining the types and sizes of machinery that could be realistically considered, recognizing that definitions of off-road machinery vary across jurisdictions. These elements are embedded within Uruguay's regulatory and institutional framework, acknowledging both enabling factors and potential bottlenecks.

A face-to-face interactive workshop was held to gather insights from key actors in the construction and related sectors. Participants included machinery suppliers, major users, and consultants from relevant industry stakeholders. The workshop, hosted at the University's facilities, followed a structured sequence of activities: introductory briefing on the project and current findings; group-based brainstorming to identify barriers and challenges; prioritization exercise using a sticker-voting system; group presentations and plenary discussion; final brainstorming round to identify opportunities and potential solutions. This process enabled the identification and prioritization of perceived challenges and facilitated the generation of preliminary opportunity pathways for electric machinery adoption.

Complementing the workshop, semi-structured interviews were conducted with key industry stakeholders selected for their operational expertise and direct involvement in procurement, fleet management, and equipment maintenance. The purposive selection of expert informants followed established qualitative research principles [20]. The interviews were carried out in person and lasted approximately forty minutes each. Their purpose was to capture perceptions of electric machinery, identify operational constraints, and

detect barriers and opportunities for adoption. Notes were taken during each interview and analysed through thematic coding to extract convergent insights, divergences, and recurring patterns.

In parallel, a structured survey was administered to characterize the existing machinery fleet, assess respondents' familiarity with electric machinery at national and international levels, and measure interest in adopting electric alternatives. The target population included suppliers, operators, fleet supervisors, and consultants engaged in construction and forestry activities. The instrument was designed and distributed digitally via SurveyMonkey to 64 identified actors. Additional dissemination was conducted through sectoral organizations such as the Uruguayan Construction Chamber, as well as through a public post on LinkedIn via the official CINOI account. Data collection took place throughout November 2024, and all responses were treated confidentially and used solely for academic purposes.

Complementing the qualitative and survey components, a technical and techno-economic assessment was undertaken to identify and evaluate an appropriate electric loader for the total cost of ownership (TCO) analysis. For the TCO assessment, a quarry operation was selected as a representative case study. The company's diesel-powered fleet was reviewed, and candidate machines for electrification were evaluated. To determine the technical feasibility of replacing the baseline equipment, a set of electric loaders was evaluated using key performance criteria, including power output, bucket capacity, maximum lift height, and breakout force. A machine was considered a viable substitute if its specifications matched or exceeded those of the reference combustion loader or achieved at least 80% of each parameter when other performance indicators remained comparable. Once the baseline equipment and the selected electrical alternatives were established, the TCO was calculated using different OPEX and CAPEX parameters as recommended by existing literature [21], [22], [23], [24], [25] & [26]. A more detailed explanation of the variables considered is included in the result section for the TCO analysis.

4. Qualitative approach: Workshop, interviews and survey results

The following section presents the integrated findings from the qualitative components of the study, which include the initial stakeholder workshop, the semi-structured interviews, and the online survey. These three sources of evidence were analysed jointly to identify convergent perceptions, recurring themes, and key insights regarding the current use of off-road machinery, the perceived barriers to decarbonization, and the feasibility of adopting alternative powertrain technologies in the Uruguayan forestry sector.

4.1 Workshop

During the workshop, participants engaged in a brainstorming exercise to identify the main barriers and challenges associated with purchasing and operating electric machinery. The ideas generated were subsequently classified into four categories—Infrastructure, Cost, Training, and Technology. The results are shown in Figure 1. A second brainstorming activity was then conducted to explore potential solutions to the previously identified challenges, as well as opportunities that could incentivize companies to adopt electric machinery. After each activity, the groups presented their outputs to facilitate discussion and promote collective reflection.

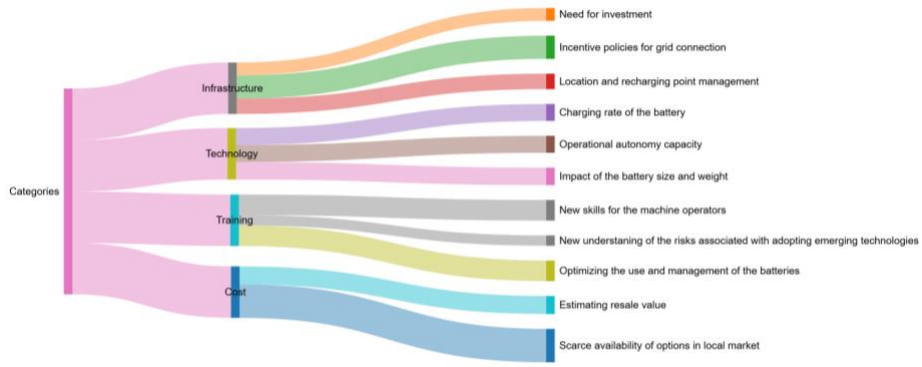


Figure 1: Sankey Diagram of the main barriers detected during the workshop

Participants identified insufficient charging infrastructure as a primary barrier to adopting electric machinery. Public chargers were considered unreliable in terms of availability, performance, and location, making operational planning difficult. Private chargers were viewed as the only practical solution, though they require substantial investment; differentiated tariffs were suggested as potential support mechanisms.

Opinions on training were mixed. While some argued that battery management requires specific instruction, others felt the machinery operates similarly to traditional equipment. Nevertheless, user adaptation—particularly managing work schedules and charging needs—was consistently identified as a challenge, especially for long or unpredictable workdays.

Cost was another major concern. Electric machines have significantly higher upfront prices and require additional charging infrastructure, although maintenance costs are lower. Limited financing options and weak resale value were also noted. A minority of participants believed that, with the right machine selection and fiscal incentives, these costs could be offset.

The discussion also highlighted a lack of information on international experiences and real-world performance, which hinders companies' ability to evaluate the financial viability of electric machinery.

Regarding opportunities, participants highlighted the need to expand and improve the public charging network, along with updating regulations and tariff schemes to provide more favorable electricity prices for charging heavy electric machinery. They emphasized several advantages of electric equipment, including lower maintenance requirements, reduced noise, and the ability to meet client demands or comply with potential future regulations. Training in the operation of electric machinery was also seen as an opportunity, as it would increase the demand for qualified operators and enhance their employment prospects.

Participants further noted that streamlining administrative processes—such as permits, approvals, and homologation procedures—would help avoid operational delays, ideally reducing processing times to less than one month. Finally, they proposed establishing a collaborative network among multiple stakeholders to develop strategically located charging points, recognizing the operational and economic benefits such infrastructure would provide.

4.2 Interviews and survey

The interviews and survey revealed several shared perspectives regarding barriers, viable applications, and future expectations for electric heavy machinery. Approximately 65% of survey respondents reported basic or intermediate knowledge of electric equipment, yet only 30% had considered incorporating it into their operations. Consistent with the workshop findings, the primary barriers identified were high upfront cost

(85%), insufficient charging infrastructure (75%), and limited battery autonomy (60%). Additionally, 40% expressed concerns about the lack of local technical expertise.

Interviewees emphasized that current charging times are excessively long—often requiring several hours—and that electrical upgrades needed to support high-power charging add cost and complexity. Companies stressed the importance of deploying faster, higher-capacity chargers, with options of 250 kW (or even 450 kW) being evaluated to shorten charging cycles to roughly one hour, thus improving feasibility for continuous operations.

Battery autonomy emerged as another major barrier, particularly in forestry and long-haul transport, where equipment operates nearly 24/7. While battery-swapping could theoretically address downtime, firms viewed this approach as impractical due to vibration-related safety concerns and the substantial infrastructure costs required.

Regulatory and administrative barriers were also highlighted, including delays of up to one year in homologation processes, high licensing costs for electric trucks, and legal restrictions that hinder rental of such vehicles. Survey results further underscored that 90% of respondents see fiscal incentives and financing mechanisms as essential for adoption, while 70% consider clear regulations and standards necessary to guide the transition.

Despite these challenges, both interviews and survey responses pointed to specific contexts in which electric machinery could be viable. These include operations with limited travel distances—such as quarries and material storage yards in industrial sites—where charging can be planned more easily, as well as urban or enclosed environments where zero emissions and low noise are strong advantages. The perceived benefits cited most frequently were reduced operating costs (70%), positive environmental impact (60%), and lower maintenance needs (50%).

5. Quantitative approach: Total cost of ownership in a quarry scenario

This section presents the main results of the case study conducted at a quarry in Uruguay. As mentioned before, electrifying off-road machinery in quarries seems particularly feasible due to several operational advantages. These sites typically have fixed infrastructure with sufficient contracted electrical capacity, allowing reliable and timely on-site charging. Equipment also operates within limited travel ranges, reducing concerns about autonomy. Moreover, standard eight-hour shifts enable charging outside working hours without affecting productivity. Together, these conditions create a favorable setting for adopting battery-electric machinery.

5.1 Case description

This study analyzes the economic performance of two alternative wheel loaders, one combustion-powered, and the other battery-powered, for use in a quarry in Montevideo. For this study, the first step was to identify the diesel equipment to be replaced. In coordination with the company operating the quarry, it was determined that the most suitable candidate for substitution was the front-end loader, hereafter referred to as Model A. To assess the technical feasibility of replacement, several electric front-end loader models were analyzed and compared. The evaluation criteria included power rating, bucket capacity, maximum lift height, and maximum breakout force. A minimum replacement threshold was established whereby the electric alternative must meet at least 80% of the performance parameters of the original diesel machine. Based on this comparative assessment, an electric front-end loader (Model B) was identified as a viable alternative. The technical characteristics of both models are presented in the table below.

Specification	Model A (Diesel)	Model B (Electric)
Power (kW)	148	240
Bucket capacity (m³)	2.7	3.5
Maximum lift height (mm)	2,999	3,500
Maximum breakout force (kN)	165	167
Operating autonomy (h)	–	6–8
Battery capacity (kWh)	–	282

Table 1 - Model specifications

To ensure that the Total Cost of Ownership (TCO) analysis remained independent of operational specifics, the calculations were performed on an annual basis. This required defining the annual operating hours of the equipment. Historical operational records provided by the company for the past five years indicate an average of 1,700 hours per year, assuming 250 operating days annually. This gives an average of 6.8 hours of operation per day, so to ensure functionality of the electric loader two charging times were considered. One during lunch break (of 45 minutes) and one full charge during nighttime.

5.2 TCO calculation and variables

For the purposes of this study, and based on existing literature [22], [24] & [26] the following formula is used to calculate the TCO:

$$TCO\ (USD) = \text{Acquisition Cost} + \sum_{n=0}^{n=10} \frac{Energy\ cost}{(1+r)^n} + \sum_{n=0}^{n=10} \frac{\text{Maintenance cost}}{(1+r)^n} - \frac{Resale\ value}{(1+r)^{10}}$$

Since r represents a discount rate, fuel costs, maintenance costs, and the resale value are treated as cash flows that occur in the future rather than at the time of equipment purchase. Assuming an annual depreciation rate of 10%, these values are discounted to their present value, meaning the value they would have if they were to be paid or received today.

The acquisition cost includes the purchase price of the equipment. For the electric machine, this also includes the cost of the charger installation. In addition, for the electric machine, the acquisition cost is reduced by the fiscal incentive corresponding to energy-efficiency projects. Specifically, the electric loader (machine + charger installation) has an acquisition cost of USD 246,000, from which USD 104,402 is discounted in fiscal incentives over a two-year period. For the diesel machine, the total acquisition cost considered is USD 91,000.

Regarding energy consumption costs, the electric machine results in an annual cost of USD 5.661, whereas the diesel machine incurs a fuel cost of USD 16,591 per year.

For maintenance, the analysis considers the costs associated with consumables (filters, oils and lubricants, coolants) and labor. Based on calculations using data provided by the company and equipment manufacturers, the annual maintenance cost amounts to USD 1,799 for the electric machine and USD 3.742 for the diesel model.

Furthermore, two cases were considered for the maintenance cost, due to uncertainty regarding the longevity of battery life in the electric model. In one case the maintenance cost included the replacement of the battery in the electric loader in year eight, which accounts for USD 96.000, and, in the case of the diesel model, an overhaul is considered in year nine, for USD 45.500. Second case does not include costs of battery replacement nor of machine overhaul.

Finally, the TCO calculation includes the resale value of the machine, which represents the estimated amount recoverable upon selling the equipment at the end of its operational life. This value acts as a future inflow and partially offset the initial investment made by the company. In this study, the resale value is assumed to be 20% of the machine's purchase price, discounted to present value using a 10% discount rate, for the case

of battery replacement and machine overhaul. In case these costs are not considered, 10% of the machine's purchase prices is considered as the resale value.

5.3 TCO main results

First, acquisition, fuel, and maintenance costs were obtained for both technologies. The results are presented in Figure 2 and 3. As can be seen, although the initial purchase cost of the electric machine is higher, the savings in energy consumption and maintenance, combined with fiscal incentives for adoption of clean technology, make the total cost over the 10-year period comparable between the two alternatives.

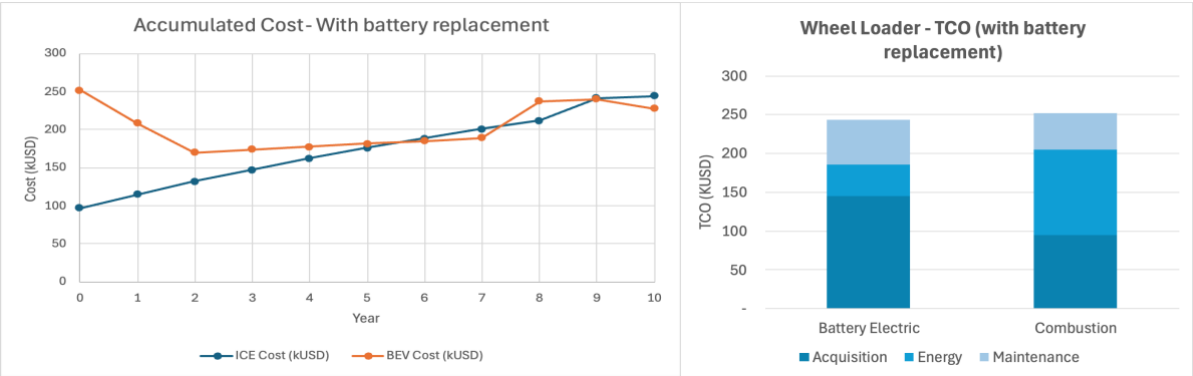


Figure 2: Accumulated TCO for both models (electric and diesel) considering battery replacement and machine overhaul

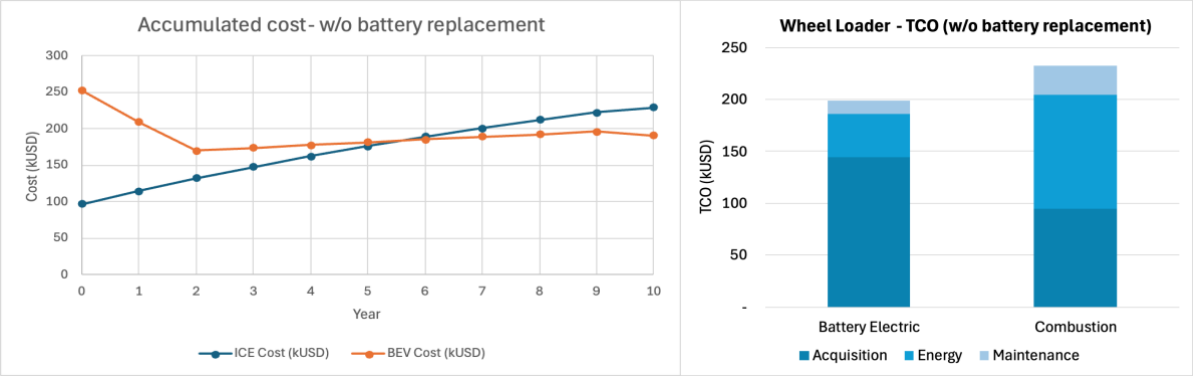


Figure 3: Accumulated TCO for both models (electric and diesel) without battery replacement and machine overhaul

In the first analyses (considering battery replacement) the cost of the electric option is 7% cheaper than its diesel counterpart. If no battery replacement is considered the electric option is 20% cheaper than its diesel counterpart.

A sensitivity analysis was also done for the battery replacement case. The analysis showed that if price of electricity increases a 13% annually (compared to the actual 4.5% annual increase, almost three times the current percentage increase) then the diesel machine will be cheaper than the electric counterpart. Furthermore, the sensitivity analysis shows how tax incentives are important for decision making processes. If the tax incentive is not considered, then the electric loader would be 30% more costly than its diesel counterpart.

In addition to the economic component, transitioning to electric machinery also reduces greenhouse gas (GHG) emissions. To quantify this impact, the "Non-Road Mobile Machinery Emissions Calculation Tool"

developed by the Climate and Clean Air in Latin American Cities Program (CALAC+) [17] was used. By inputting key assumptions—250 working days per year, a 10-year study period, diesel sulphur content in ppm (50 ppm for Gasoil 50-S), and 1,700 annual operating hours—the tool calculates GHG emissions over the selected period. The results allow estimation of the emissions that would be generated if the front-end loader with a combustion engine were used during the 10-year study. Based on the data from this study, the electric loader avoids the emission of 47.9 kg of CO_2 per hour of operation, or 814 tons of CO_2 across the 10-year study.

6. Conclusions and future research lines

This study assessed the feasibility of adopting battery-electric machinery in Uruguay's heavy off-road sector, focusing on construction and quarrying. Through literature review, stakeholder workshops, interviews, surveys, and a Total Cost of Ownership (TCO) analysis, the research identified key technical, economic, and operational considerations for transitioning from diesel to electric equipment. Results indicate that although electric machinery has higher upfront costs, savings in energy, maintenance, and fiscal incentives make the total 10-year cost comparable to diesel alternatives, while also having environmental benefits by avoiding CO_2 emissions during its operation. Quarries and similar sites are particularly suitable for early adoption due to fixed infrastructure, high power availability, limited equipment travel, and defined working hours that allow off shift charging. The study also highlighted operational barriers including limited technical expertise, the need for reliable charging infrastructure, and regulatory challenges. Opportunities to support adoption include expanding public and private charging networks, streamlining permitting processes, offering targeted training programs, and implementing fiscal or financial incentives. Stakeholder collaboration was highlighted as key to bridging information gaps and accelerating adoption.

Future research should expand the scope to include other types of off-road machinery, such as hybrid or hydrogen-based equipment, once technological maturity and local supply chains improve. Additional studies could also analyse fleet-level adoption, integration with renewable energy sources, battery lifecycle management, and long-term operational performance in different industrial contexts. Collecting real-world data from early adopters in Uruguay and across Latin America will be critical to refine economic and environmental models and guide policy development.

In conclusion, this study demonstrates that battery-electric heavy machinery is a technically and economically feasible option for specific industrial applications in Uruguay, with clear potential for environmental and operational benefits. With coordinated policy, infrastructure investment, and stakeholder engagement, electromobility in the off-road sector can become a viable pathway toward a more sustainable industrial landscape.

Acknowledgements

This study forms part of the FSE_S_2023_1_179284 project, funded by the National Agency for Research and Innovation in Uruguay (ANII).

References

[1] S. Amin. 2025. Decarbonisation of Road Transportation in India—A Round-Robin Review on Low-Carbon Strategies and Financial Policies, Future Transportation, vol. 5, n.o 2, p. 36, doi: 10.3390/futuretransp5020036.

[2] U. Motowidlak. 2022. Conditions for the sustainable development of electromobility in the European Union road transport from the perspective of the European Green Deal», Prace Komisji Geografii Komunikacji PTG, vol. 2022, n.o 25 (4), pp. 7-25.

[3] A. I. NUR, H. HERYANA, y A. R. NUR, «EXAMINING THE CONTRIBUTION OF LOCAL GOVERNMENT IN ADVANCING ELECTROMOBILITY: INSIGHTS FROM THE LEGAL FRAMEWORK IN INDONESIA.», Juridical Current, vol. 27, n.o 2, pp. 43-60, abr. 2024, doi: 10.62838/CJJC.97.2.11.

[4] British Petroleum (BP). 2023. «BP Energy Outlook».

[5] MIEM, DNE. 2023. Balance Energético 2023.

[6] MIEM, DNETN. 2005 Política Energética 2005-2030.

[7] «Proyecto MOVÉS», Ministerio de Industria, Energía y Minería. Accedido: 18 de julio de 2025. [Online. available at: https://www.gub.uy/ministerio-industria-energia-mineria/proyecto-moves

[8] Ministerio del Medio Ambiente Chile. 2025. ¿Qué son las fuentes móviles de emisiones contaminantes? – Fuentes Móviles. Accedido: 23 de julio de 2025. [En línea]. available at: https://fuentesmoviles.mma.gob.cl/que-son-las-fuentes-moviles-de-emisiones-contaminantes/

[9] Ministerio del Ambiente Perú. 2022. Inventario de emisiones de maquinaria móvil no de carretera (MMNC) en el Perú y en Lima Metropolitana y Callao, Ministerio del Ambiente Perú.

[10] F. E. G. Mendes, D. I. Brandao, T. Maia, y J. C. Braz de Filho. 2019. Off-Road Vehicle Hybridization Methodology Applied to a Tractor Backhoe Loader, en 2019 IEEE Transportation Electrification Conference and Expo (ITEC), pp. 1-6. doi: 10.1109/ITEC.2019.8790624.

[11] D. Troncon, L. Alberti, S. Bolognani, F. Bettella, y A. Gatto. 2019. Electrification of agricultural machinery: a feasibility evaluation, Fourteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), may 2019, pp. 1-7. doi: 10.1109/EVER.2019.8813518.

[12] S. Wieser, L. Brünner, M. Halfen, F. Heckert, y M. Böhm. 2024. Hydrogen in Off-Road Vehicle Applications, Third International Conference on Sustainable Mobility Applications, Renewables and Technology (SMART), nov. 2024, pp. 1-6. doi: 10.1109/SMART63170.2024.10815319.

[13] E. E. Pompodakis, A. Ahmed, C. Condaxakis, G. I. Orfanoudakis, y E. S. Karapidakis. 2024. Paving the way toward hydrogen mobility: An effective policy to reduce the green hydrogen cost, International Journal of Hydrogen Energy, vol. 100, pp. 417-427, doi: 10.1016/j.ijhydene.2024.12.073.

[14] I. Krzyżewska y K. Chruzik. 2023. Maintenance and Exploitation of Electric, Hybrid, and Internal Combustion Vehicles, Energies, vol. 16, n.o 23, p. 7842, doi: 10.3390/en16237842.

[15] F. Kuipers y R. Wolbertus. 2025 Smart charging of battery-electric construction machinery at construction sites; A case study in the municipality of The Hague.: 38th International Electric Vehicle Symposium and Exhibition, Disponible en: https://evs38-program.org/en/

[16] A. HUSSAIN. 2025. The Electric Revolution in Heavy Mobile Machinery Manufacturing, Dassault Systèmes blog. Accedido: 12 de diciembre de 2025. [En línea]. Disponible en: https://blog.3ds.com/industries/industrial-equipment/the-electric-revolution-in-heavy-mobile-machinery-manufacturing-2/

[17] Hagan, R., Markey, E., Clancy, J., Keating, M., Donnelly, A., O'Connor, D. J., Morrison, L., & McGillicuddy, E. J. 2023. Non-Road Mobile Machinery Emissions and Regulations: A Review. Air, 1(1), 14-36. https://doi.org/10.3390/air1010002.

[18] M. Šušnjar, Z. Pandur, M. Bačić, K. Lepoglavec, H. Nevečerel, y H. Kopseak. 2023 Possibilities for the Development of an Electric Hybrid Skidder Based on Energy Consumption Measurement in Real Terrain Conditions, Forests, vol. 14, n.o 1, p. 58, doi: 10.3390/f14010058.

[19] Regal. 2025. Electric Forestry and Agricultural Machinery Regal. Access online: 12/12/2025. Available at: https://regal.se/insights/electric-forestry-and-agricultural-machinery/

[20] Creswell, J.W. and Creswell, J.D. (2018) Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage, Los Angeles.

[21] C. Rout, H. Li, V. Dupont, y Z. Wadud. 2022. A comparative total cost of ownership analysis of heavy duty on-road and off-road vehicles powered by hydrogen, electricity, and diesel, Heliyon, vol. 8, n.o 12, doi: 10.1016/j.heliyon.2022.e12417.

[22] B. Noll, S. del Val, T. S. Schmidt, y B. Steffen. 2022. Analyzing the competitiveness of low-carbon drive-technologies in road-freight: A total cost of ownership analysis in Europe, Applied Energy, vol. 306, doi: 10.1016/j.apenergy.2021.118079.

[23] D. Rajagopal, N. Gopinathan, A. Khandekar, N. Karali, A. Phadke, y N. Abhyankar. 2024. Comparative Evaluation of Total Cost of Ownership of Battery-Electric and Diesel Trucks in India, Transportation Research Record, vol. 2678, n.o 6, pp. 235-247, doi: 10.1177/03611981231195055.

[24] Martín Tanco, Luis Cat, Santiago Garat. 2019. A break-even analysis for battery electric trucks in Latin America, Journal of Cleaner Production, Volume 228, Pages 1354-1367, https://doi.org/10.1016/j.jclepro.2019.04.168.

[25] Qiu, Yin, Cristina Dobbelaere, and Shuhan Song. 2023. "Energy Cost Analysis and Operational Range Prediction Based on Medium- and Heavy-Duty Electric Vehicle Real-World Deployments across the United States" World Electric Vehicle Journal 14, no. 12: 330. https://doi.org/10.3390/wevj14120330

[26] M. A. Saafi, V. Gordillo, O. Alharbi, y M. Mitschler. 2024. Investigating the Future of Freight Transport Low Carbon Technologies Market Acceptance across Different Regions, Energies, vol. 17, n.o 19, p. 4925, doi: 10.3390/en17194925.

Biography

Daniel Jurburg - Received the degree in industrial engineering from Universidad de Montevideo, Uruguay, the master's degree in applied engineering from University of Navarra, Spain, and the Ph.D. degree in engineering from the Applied Engineering Program, TECNUN—University of Navarra. He is currently the Director of the Center for Innovation in Industrial Management (CINOI), School of Engineering, Universidad de Montevideo. He participated in several research projects and has published several articles in indexed international journals on supply chain management and continuous process improvement topics.

Marcelo Berglavaz – Industrial Engineer specialized in energy systems. He currently serves as the Industrial Engineering Program Coordinator and as a Researcher at CINOI and CPML at the Universidad de Montevideo. He also works as a mechanical and industrial engineering consultant for Matriz. His professional experience spans energy efficiency, hydrogen technologies, renewable energy—particularly biomass—and the oil sector. He is currently pursuing a Master's degree in Applied Research in Engineering, focusing on advanced energy and industrial applications.

Justino Romano – Industrial Engineer specialized in Industrial Organization and Production from Universidad de Montevideo, with experience in industrial design, logistics, maintenance, and environmental projects. He is currently pursuing a Master's degree in Applied Research in Engineering. He serves as an Operations Functional Consultant at Dynamo Global and as a Researcher at the School of Engineering, Universidad de Montevideo.

Rodrigo Sanchez - He holds a degree in Industrial Engineering from Universidad de Montevideo and holds a postgraduate degree in Data Science and Big Data. Since 2019, he has participated in several research and applied engineering projects at CINOI, focusing on transportation and logistics. Additionally, he has experience in Data Analytics applied projects in different industries. Currently, he is a researcher and lecturer at the School of Engineering of Universidad de Montevideo.

Camila Perez – Advanced Industrial Engineering student at Universidad de Montevideo. She is currently part of a research team focused on electromobility in traditional sectors such as construction and forestry. Throughout her studies, she has demonstrated a strong interest in applied technological innovation.

Micaela Montilla – Advanced Industrial Engineering student at Universidad de Montevideo. Since 2024, she has participated in research and applied engineering projects at CINOI, focusing on heavy electric mobility. She also has experience in hydraulics projects. Currently, she works as a researcher and teaching assistant at the School of Engineering of Universidad de Montevideo.

8th Conference on Production Systems and Logistics

Implementation of a Digital Twin Prototype For The After-Sales Service Workshop Process

Afonso Silva[1], Paulo Ávila[1,2], João Bastos[1,2], Luís Pinto Ferreira[1,3], Hélio Castro[1,2], Alzira Mota[1,2]

[1]ISEP, Polytechnic of Porto, rua Dr. António Bernardino de Almeida, Porto, Portugal
[2]INESC TEC - Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal
[3]Associate Laboratory for Energy, Transports and Aerospace (LAETA-INEGI), rua Dr. Roberto Frias 400, Porto, Portugal

Abstract

This article outlines the development and feasibility study of implementing a Digital Twin (DT) prototype in the after-sales service management system of Company X, a truck and bus dealer in Portugal. The objective of the project is to contribute to the digitization of the service workshop process, and in the future enabling real-time monitoring and control through sensors and an integrated platform. The research commenced with the detailed modelling of the process using BPMN (Business Process Modeling Notation) in Bizagi software, followed by the construction of a digital model in FlexSim software. This model was parameterised with real data, and a dashboard was developed to monitor the process. Furthermore, a connection has been established between the digital model and a Python algorithm, with a view to the efficient management of intervention stations. This paper presents the design and implementation of a Digital Twin–inspired prototype (currently at a Digital Shadow level) for an after-sales service workshop. The contribution is a process-to-simulation pipeline, KPI monitoring dashboard, and exploratory data-driven analysis of workshop operations based on seven days of real data.

Keywords

Digital Twin Prototype; Service Workshop Process; Operational Efficiency.

1. Introduction

In the automotive sector, effective after-sales service is critical to customer satisfaction and loyalty. However, this is often hindered by challenges related to operational inefficiency and a lack of accurate data for strategic decisions [1]. Research indicates that after-sales services can be a valuable source of revenue and profit for companies. In the industrial sector, for instance, after-sales services have been found to contribute up to 25% of total revenue and 40–50% of profits [2]. In this context, the implementation of Digital Twins emerges as an innovative solution to optimise processes, improve key performance indicators (KPIs) and increase competitiveness in the sector. In recent years, the concept of Digital Twin has gained prominence in the industrial sector, enabling the creation of digital representations of physical systems that are used to monitor, simulate, and optimise operations. The term was initially proposed by Grieves [3] in the context of the product life cycle, and since then, its application has expanded to various areas, including manufacturing and healthcare. The potential of Digital Twins to increase operational efficiency and reduce costs is especially relevant in complex and dynamic processes, such as after-sales services. The digitisation

DOI: https://doi.org/10.15488/21028
ISSN: 2701-6277

of workshop processes in after-sales service is a growing necessity, given the importance of monitoring Key Performance Indicators (KPIs) accurately and in real time. However, in many cases, the data required for this monitoring is collected manually, which can result in errors and delays. At Company X's after-sales unit, this limitation creates challenges in identifying inefficiencies and making data-driven decisions. The primary objective of this study is to develop and validate a Digital Twin prototype to monitor and explore optimisation opportunities of after-sales processes, with a focus on:

- Analyse and document the workshop process of Company X's after-sales service;
- Specify the technical and functional requirements of a Digital Twin prototype;
- Develop and validate a Digital Twin prototype.

This study is guided by two main research questions. First, it investigates how a Digital Twin–inspired prototype can be designed and implemented to support the monitoring of Key Performance Indicators (KPIs) in an after-sales service workshop. Second, it examines the technical and functional requirements that such a prototype should satisfy in order to enable future integration with data collection sensors and decision-support or control components.

Although the use of Digital Twins has been widely studied in industrial environments [4], there is a gap in the application of this technology in after-sales workshops, where processes are highly dependent on human intervention. The research will contribute to the field of Digital Twins by demonstrating the practical application of this concept in an after-sales service context, addressing existing barriers and proposing technological solutions to them. The article is structured as follows: Section 2 presents a literature review, highlighting related works and the state of the art in digital twins and process digitisation. Section 3 provides a comprehensive overview of the practical work, encompassing the implementation process and the results obtained. Finally, section 4 presents the conclusions and suggestions for future work.

2. Literature Review

The concept of a digital twin, initially proposed by Grieves in 2002 [5], emerged in the context of product lifecycle management and quickly expanded to various industries due to its potential to optimise operations and reduce costs. In the context of digital twin research, the works of Putnik et al. [6], Ávila et al. [7] and Cunha et al. [8] can be framed as an early and complementary contribution to enterprise-level and decision-oriented digital twins. Their researches focus on modelling and coordinating agile and virtual enterprises, and selecting resources in agile and virtual enterprises with the support of a broker and the market of resources, which aligns with the core objectives of digital twins: representing complex systems digitally to support analysis, decision-making, and improved performance. Although their works does not explicitly adopt the digital twin terminology, it addresses fundamental concepts that underpin organizational digital twins. Digital twins are digital replicas of physical processes, systems, or products that enable performance monitoring, simulation, and analysis in real time. Its current applications extend to areas ranging from production and logistics to healthcare and equipment maintenance. In recent literature, several studies have explored the concept of Digital Twins to monitor and optimise processes. For instance, the integration of Digital Twins with IoT sensors has been analysed in order to facilitate real-time monitoring of industrial processes, emphasising the role of this technology in enhancing data accuracy and accelerating responses to operational issues [9]. A relevant study is that of Kritzinger et al. [4], which presents a systematic review of the literature on Digital Twins in the context of Industry 4.0, with a focus on production. This study categorises Digital Twins into three levels of integration: The Digital Model, Digital Shadow and Digital Twin are presented, with particular emphasis on the latter as the most advanced level, which incorporates real-time bidirectional integration between the digital model and the physical system. Despite the growing use of Digital Twins, the authors note that the majority of publications continue to concentrate on theoretical concepts and initial studies, with practical applications such as predictive maintenance and adaptive planning

remaining in the early stages. Eyring et al. [10] developed digital twins to monitor and optimise the layout and workflow on a production line using FlexSim software. This study concluded that the use of digital twins enabled a significant reduction in waiting times and workflow inefficiencies. Coleman and Proctor (2021) [11] developed a digital twin for the glass industry with the aim of optimising warehouse layout and improving space management. This study highlighted the benefits of simulation for analysing different layout scenarios and making decisions based on accurate data. Despite these advances, the literature lacks studies applied to after-sales workshops, particularly in the development of Digital Twins adapted to continuous service processes and performed largely by human operators. This study seeks to fill this gap by studying the implementation of a Digital Twin prototype in an after-sales unit with sensor integration and computer simulation. The main contribution of this study is to demonstrate how Digital Twins can be applied in a practical and effective way to monitor and control performance indicators in a workshop, providing valuable insights for other companies in the sector.

In the present work, the implemented system corresponds more closely to a Digital Shadow or a transitional Digital Twin, since bidirectional real-time control is not yet deployed. The prototype focuses on monitoring and decision-support capabilities based on data-driven simulation and therefore represents a first step towards a more comprehensive Digital Twin architecture.

3. Practical work

One of Company X's key offerings is its after-sales service, which is provided to customers of a particular brand of lorries and buses in Portugal. At present, the company is experiencing difficulties in obtaining the data necessary to calculate the KPIs for the after-sales service workshop process. This is due to the fact that the execution of the process is mostly the responsibility of humans and not machines. Therefore, given that it is neither feasible nor efficient to rely on manual collection of the information needed for the indicators, the company has created a project to digitise the process. The aim of the project is to develop a Digital Twin that will allow the workshop's activities to be viewed and controlled digitally.

3.1 Methodology

This study follows a design-oriented case study methodology aimed at developing and assessing a Digital Twin–inspired prototype for an after-sales service workshop. The research focuses on the design, implementation, and exploratory evaluation of a digital representation of a real workshop process, with the objective of assessing technical feasibility and the ability of the prototype to support operational monitoring through Key Performance Indicators (KPIs).

The methodology is structured into five main phases. First, the after-sales service process was analysed and modelled using Business Process Model and Notation (BPMN) in order to obtain a structured and shared representation of the real workflow. This phase aimed to clarify process logic, identify key activities, and define the system boundaries for the digital model.

Second, a discrete-event simulation model of the workshop was developed in FlexSim, representing the main resources, queues, and process stages observed in the real system. The digital model was built to reflect the physical layout and operational logic of the workshop and to enable the collection of performance indicators through a dedicated dashboard.

Third, real operational data were collected over a period of seven working days and used to parameterise the simulation model. The collected data were analysed and fitted to statistical distributions using ExpertFit in order to represent the current state of the process in a consistent and reproducible manner. This step aimed to improve the realism of the simulation while acknowledging the limited size of the available dataset.

Fourth, the parameterised simulation model was executed to perform an exploratory evaluation of the system behaviour and to generate KPIs related to waiting times, service durations, and resource utilisation. The purpose of this evaluation is not to provide statistically generalisable performance claims, but rather to assess

the feasibility of the approach and to demonstrate the type of operational insights that such a digital prototype can provide.

Finally, a connectivity component was implemented by integrating the FlexSim model with an external Python-based decision-support algorithm. This integration illustrates how the digital model can exchange data with external logic to support operational decisions, thereby representing a transitional step towards a more advanced Digital Twin architecture.

Overall, the proposed methodology is exploratory and feasibility-oriented. Given the limited data availability and the proof-of-concept nature of the implementation, the results should be interpreted as illustrative rather than conclusive. The study is intended to demonstrate the technical viability of the approach and to provide a foundation for future extensions, such as real-time data integration, bidirectional control, and predictive or optimisation-based functionalities.

3.2 Analysis of the after-sales service process

The after-sales service process (Figure 1) is a series of steps that begin when the customer's lorry arrives at reception, continue with work on the vehicle, and conclude with collection by the customer.

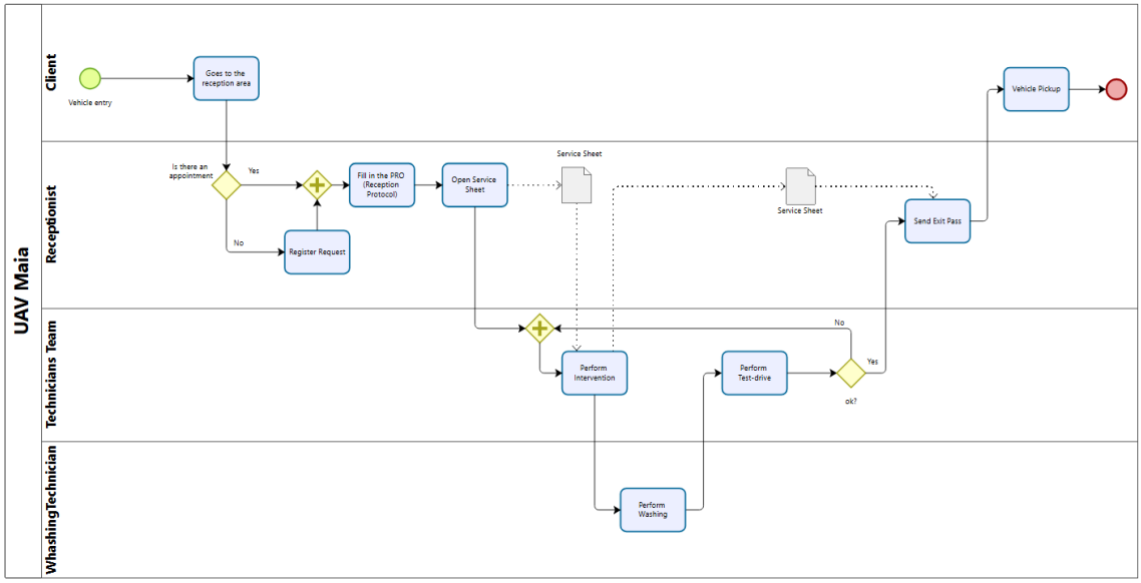


Figure 1: After-sales process diagram

The process begins when the customer arrives at the after-sales unit and goes to reception. The receptionist then checks whether there is a service appointment and carries out the Reception Protocol (PRO), whose main purpose is to collect the information needed to start the service, such as mileage, vehicle condition and customer signature. Once the PRO is complete, the car is parked in front of the workshop, ready to be serviced. When the order is given, the vehicle enters the workshop, to a station, for work to begin. Depending on the number, the stations may have the capacity for one or two lorries. The stations with greater capacity are primarily intended for buses due to their greater length compared to lorries. After the vehicle has been serviced, it is placed in a queue for washing. Washing consists of cleaning the entire interior of the cabin, as well as the exterior of the vehicle. Next, a road test is carried out, which consists of taking the car outside the premises and checking that it is working properly. If everything is in order, the vehicle is placed in the entrance area of the premises to be picked up. Otherwise, the necessary adjustments are made in the workshop and the procedure follows the steps of washing and road testing again.

3.3 Vision and Digitisation Project

The final vision for implementing the actual system is described in the Digital Twin architecture (see Figure 2). The reader sensors would capture vehicle number plates, recording entries at the various stages of the

process. This would establish a connection with the company's MES management system database, enabling real-time data transfer. The company's management system would be directly connected to a broker that serves as an intermediary agent, enabling communication between the different environments of the Digital Twin model in FlexSim and its respective database. The digital model would enable real-time monitoring and control of the process.

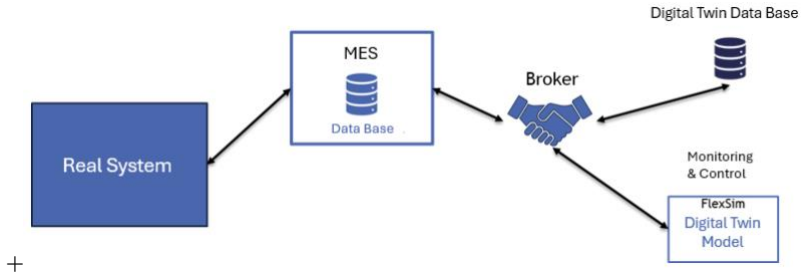


Figure 2: Digital Twin Architecture

The project phase diagram is shown in Figure 3, representing the steps to be followed for the development of the Digital Twin prototype in the context presented in Introduction. The left-hand side of the Figure 3illustrates the steps for creating and validating a functional digital model that is as similar as possible to the actual process. The phases on the right provide the digital model with the capability to analyse and intervene in the process, thus satisfying the necessary conditions for being considered a DT.

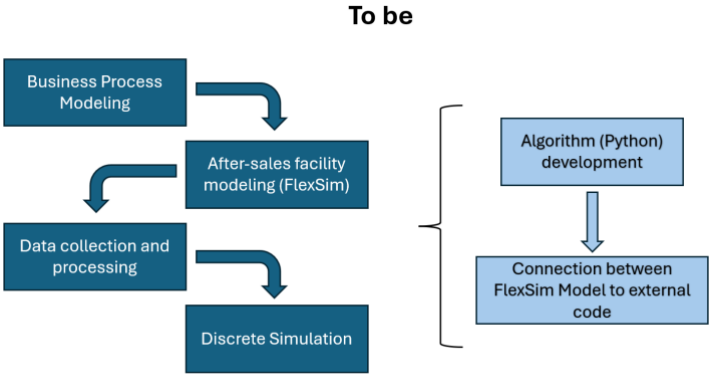


Figure 3: Project phases

The first phase involves modelling the after-sales business process and aims to enable a digital representation of the system. The next step is to model the after-sales unit on the digital platform. This will incorporate the resources and synergies inherent to the process. To validate the previously defined digital model, data is collected in the workshop and processed by adjusting statistical distributions. The testing and validation of the model is concluded with discrete simulation in FlexSim, following the parameterisation inputs from the previous phase. The left-hand side steps commence with the development of an algorithm in the Python language, which involves the receipt of information from the system and the subsequent return of an order or recommendation for operational management. At the end the connection between the digital platform and the external Python code is established. In this final phase, the data to be sent to the algorithm and its format are also defined.

3.4 Construction of the digital model of the after-sales unit

The subsequent stage in the digitisation process is to use FlexSim (see Figure 4) to represent the after-sales unit under analysis. At this stage, a dashboard was also developed in the same software to monitor the process. Therefore, software resources such as the queue, processor, source and sink were utilised. It is

important to note that the actual directions and measurements of the unit were taken into account during the modelling process.

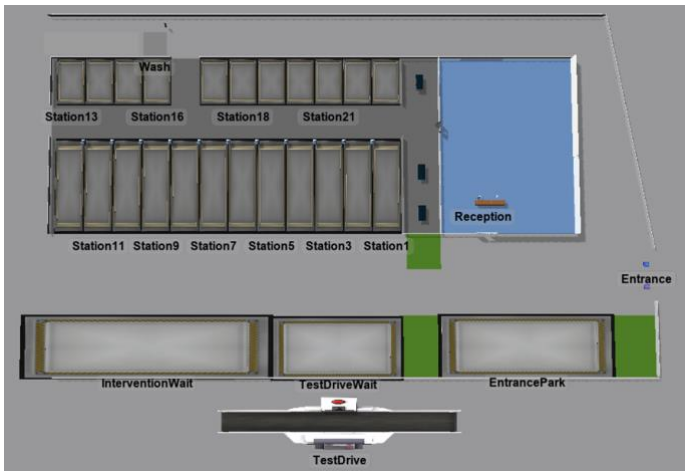


Figure 4: Digital model of the After-sales Unit in FlexSim

In the reception activity, an 'Entrance' source was added, which allows the creation of a vehicle in the digital system, a "Reception" processor that simulates the vehicle reception process, and an 'EntrancePark' queue that serves as a waiting line for reception (see Figure 5).

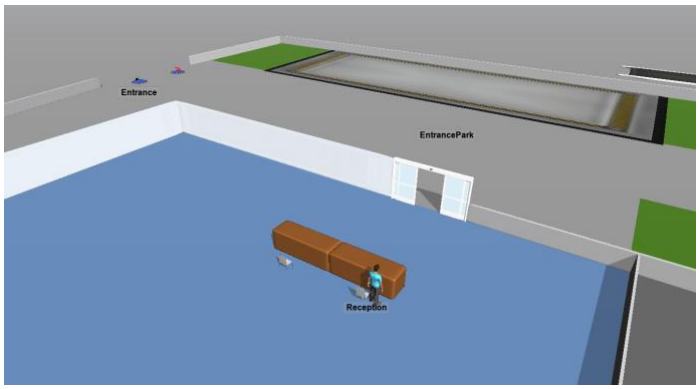


Figure 5: Reception of the digital model

During the intervention phase, a queue called 'WaitingIntervention' was used to represent the waiting area before entering the workshop. The model was expanded to include twenty-three additional processors, representing the stalls within the facilities (see Figure 6).

Figure 6: Intervention phase in the digital model

In the washing activity, a queue called 'WaitWashing' was used to represent the waiting queue for the process, and a processor called 'Washing' was added (see Figure 7).

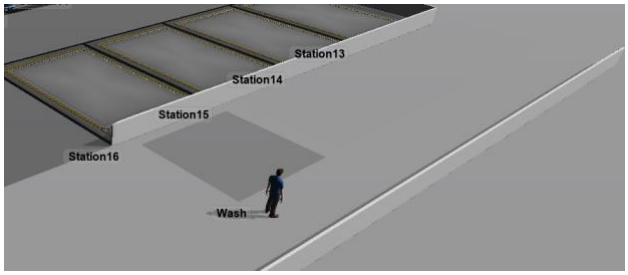


Figure 7: Washing phase in the digital model

In the subsequent phase of the process, a queue called 'WaitRoadTest' was incorporated, representing the wait until the activity in question commences. A processor called ‘TestDrive’ was also added (see Figure 8), which simulates the process referred to in its name, as described in section 3.2.

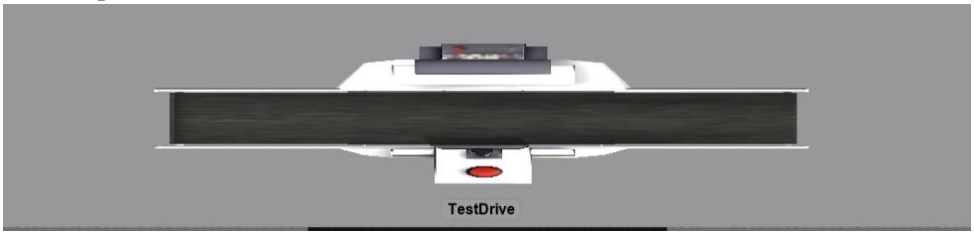


Figure 8: Road testing phase on the digital model

Additionally, FlexSim facilitates the creation of customised dashboards with graphs and tables to monitor processes. The creation of widgets was to measure waiting times at each stage of the service, such as reception, intervention, washing and road testing, with detailed tables and graphs for quick analysis. These tools offer precision and efficiency in controlling the workshop process.

3.5 Digital Model Parameterisation

In order to represent the current state of the process in the FlexSim digital model, data was collected over a total of seven days, from 8 a.m. to 6:30 p.m. The collection dates were chosen to ensure maximum continuity of the process and to allow each vehicle to be tracked throughout all stages of the process. Another objective of the data collection was to calculate operational KPIs that the company did not have full visibility of, such as: order opening duration, waiting time to enter the workshop, service duration in the workshop, washing waiting time, washing duration, road test waiting time, duration of road test, service. The collection methodology was based on manual recording on a pre-printed grid, and the data was quantitative in nature, except for vehicle registration numbers. The total results of the collection are shown in Table 1. The number of observations differs between KPIs due to the variation in duration between process activities and some lack of continuity in data collection. In addition, the relatively low number of records for some indicators may influence the accuracy of the sample.

Table 1: Data collection results

	Check-in Duration	Wait time until entering facility	Intervention time	Wash Wait time	Wash time	Test-drive wait time	Test-drive	Service time
Observations	89	62	43	26	43	23	48	33
Average Staytime (h)	0.65	5.66	5.98	1.48	0.83	3.75	0.74	22.41

The percentage of vehicles that, after the road test, return to the facility instead of completing the service was also calculated, and this figure was 8.33%.

3.6 Simulation of the Current State

At this stage, the objective was to continue the testing and validation process of the model by entering the collected data into the simulator. In order to digitally represent the current state of the process, it was necessary to change the parameters of the simulator's various resources, such as the interval between vehicle arrivals or reception time. In order to ensure the highest possible accuracy and realism in the simulation, the set of records was adjusted to statistical distributions using ExpertFit (see Table 2). This solution enables the current state to be represented with greater reliability, despite the number of observations falling short of the desired threshold.

Table 2: Distributions and their parameters adjusted to the data

Data	Distribution	Parameters
Frequency of vehicle arrivals	Exponential	$\lambda = 1.3285$ γ=-1,0 $* 10^{-14}$
Vehicle Check-in Time	Log-logistic	α =2.5414 β =0.55075 γ = -0.05071
Intervention time	Pareto type 2	α=2.5762 β= 9.8462
Whash time	Pearson type 6	$\alpha_1 = 22.081$ $\alpha_2 = 12.467$ β= 0.59233 γ = -0.27057
Test-drive time	Log-logistic	α = 2.7062 β= 0.44907 γ= 0.1314

The simulation was planned to run for seven days with a warm-up time of 23 hours so that the system would not start without any vehicles. The average waiting times per park are shown in Table 3. The average waiting time for the service was always zero because in the simulation the resources for interventions are infinite; therefore, there are always stations available. The time for “ParkEntry” corresponds to the collection field “Order Opening Duration” and obtained a value of 1.31 hours, which was slightly higher than that of the actual sample. Waiting times for washing and road testing were lower in the simulation and may be due to the occurrence of breaks and other random events in real life.

Table 3: Average waiting time per car park

Station	Avg. time(h)
EntrancePark	1.31
ServiceWait	0
TestDriveWait	0.62
WashWait	0.56

The intervention time was also calculated taking into account the 23 stations (see table 4), yielding an average duration of 15.82 hours, a value higher than the data collection, which can be explained by the greater continuity of recording of the digital model in relation to the actual sample. As for the remaining two phases, the values remained similar to the data collection, noting that there was a road test that reached 4.53 hours (see Table 5).

Table 4: Intervention time of the set of stations at the end of the simulation

Avg. Time	Min. Time	Max. Time
15.82	9.93	47.36

Table 5: Duration of washing and road-testing phases after simulation

Phase	Avg. Time	Min. Time	Max. Time
Wash	0.76	0.17	1.75
Test-drive	0.77	0.16	4.53

3.7 Evaluation and Performance Analysis

3.7.1 Identified Bottlenecks and Operational Insights

The exploratory simulation revealed several operational characteristics of the current workshop process. The average service time of 22.41 hours (Table 1) indicates significant time expenditure, with intervention time averaging 15.82 hours in simulation (Table 4) compared to 5.98 hours in the limited field sample. This discrepancy likely reflects the continuous recording capability of the digital model versus the intermittent manual collection, highlighting one of the key benefits of digitization: complete process visibility.

Table 3 indicates that the simulated system produces shorter queues than those observed in the workshop data. In the simulation, the average waiting time in the entrance parking area is 1.31 hours, whereas the manually collected data suggests an average of 0.65 hours for the corresponding stage. For the later stages, the simulation reports average waiting times of 0.56 hours for washing and 0.62 hours for road testing, compared with 1.48 hours and 3.75 hours, respectively, in the field measurements.

This gap - especially for washing and road testing - suggests that the current parameterised model may not yet capture sources of delay that occur in real operations. In practice, interruptions such as staff breaks, temporary equipment unavailability, unplanned rework, and other ad-hoc tasks can accumulate and extend queues, but these effects may be simplified or absent in the current simulation setup. As a result, the discrepancy can be used as a guide for model refinement, for example by explicitly modelling disruptions (downtimes, shift schedules, priority changes, or re-entries) and re-validating the model against observed waiting times.

3.7.2 Limitations of the Current Evaluation

It is important to acknowledge the exploratory and preliminary nature of this evaluation:

1. **Limited data scope:** Seven days of manual collection with sample sizes ranging from 23 to 89 observations per KPI (Table 1) provides initial calibration but cannot capture seasonal variations, atypical demand patterns, or rare events.
2. **Simulation assumptions:** The model assumes infinite intervention capacity (zero waiting for service stations), which differs from real constraints and may underestimate actual bottlenecks during peak demand.
3. **No baseline comparison:** Without a pre-digitization baseline or alternative scenarios (e.g., different dispatching rules), quantitative performance improvement claims cannot be validated.
4. **Proof-of-concept focus:** The primary objective was to demonstrate technical feasibility and establish the digital infrastructure, not to conduct a comprehensive performance optimization study.

3.8 Connectivity Component

In order to proceed to the final phase of the project, FlexSim's external code connection features were used to test the Digital Twin concept. This involves the digital component analysing the data received by the physical system and communicating a decision. In this project, a Python algorithm was used to determine the next vehicle's bay allocation during the intervention process (see Figure 9). The decision is made based on the occupancy of each station and its average intervention time.

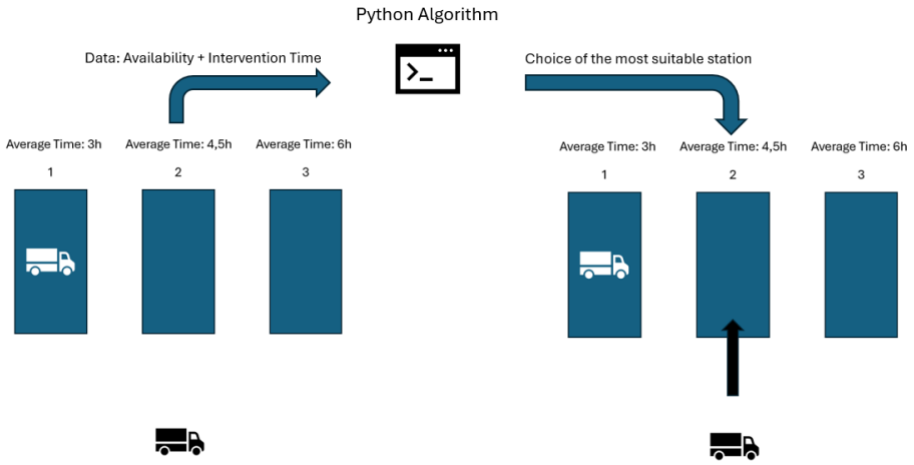


Figure 9: Example of operation with the introduction of the algorithm

4. Conclusions

This study addressed the design and implementation of a Digital Twin–inspired prototype for an after-sales service workshop, with the goal of supporting KPI monitoring and exploratory operational analysis. The results show that it is feasible to model the workshop process using BPMN and discrete-event simulation, to parameterise the model with real operational data, and to integrate it with a dashboard and an external Python-based decision-support component. In this sense, the work answers the first research question by demonstrating a practical and implementable approach for building a prototype capable of supporting KPI monitoring in an after-sales context.

Regarding the second research question, the study also highlights a set of technical and functional requirements for future developments, including the need for reliable data acquisition mechanisms, standardised data interfaces between information systems and the simulation model, and modular integration of decision-support or optimisation components. The current prototype, positioned at a Digital Shadow or transitional Digital Twin level, provides a technical foundation for future extensions such as sensor-based data collection, real-time data synchronisation, and more advanced analytics or optimisation functionalities.

Furthermore, future integration with inventory systems, predictive algorithms, and artificial intelligence could significantly expand the impact of DT in other areas of the workshop process. Despite the advances made, there are opportunities for improvements and expansions of the developed prototype. The future installation of sensors in the workshop will require adjustments to the connection to the digital model, depending on the characteristics of the sensors used. Additionally, the introduction of more advanced algorithms could enhance not only the entry into the workshop but also the subsequent stages of the process, aligning the system with the company's operational needs. The dashboard could be enhanced with new KPIs and more detailed analyses, which would provide richer data for strategic decisions. This study underlines the transformative role of Digital Twins in the digitisation of the service sector, contributing to the adoption of Industry 4.0 practices. The prototype developed represents a significant advance for Company X, improving its operational responsiveness and directly benefiting its customers. With the suggested improvements and future expansions, DT could become even more comprehensive, consolidating its position as an indispensable tool for the efficient management of the after-sales process and other operational domains.

References

[1] C. M. Durugbo, "After-sales services and aftermarket support: a systematic review, theory and future research directions," Int J Prod Res, vol. 58, pp. 1–36, Dec. 2019, doi: 10.1080/00207543.2019.1693655.

[2] J. Holmström, N. Cheikhrouhou, G. Farine, and K. Främling, "Product Centric Organization of After-Sales Supply Chain Planning and Control," in Behavioral Operations in Planning and Scheduling, 2011, pp. 187–198. doi: 10.1007/978-3-642-13382-4_9.

[3] M. Grieves, SME Management Forum Completing the Cycle: Using PLM Information in the Sales and Service Functions. 2002.

[4] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital Twin in manufacturing: A categorical literature review and classification," Ifac-PapersOnline, vol. 51, no. 11, pp. 1016–1022, 2018.

[5] M. Grieves, "Digital Twins: Past, Present, and Future," Cham, Switzerland: Springer, 2023, pp. 97–121. doi: 10.1007/978-3-031-21343-4.

[6] Putnik, G., Cunha, M., Sousa, R., Ávila, P (2005) BM Virtual Enterprise: A Model for Dynamics and Virtuality, Putnik, G. & Cunha, M. (Eds.) Virtual Enterprise Integration: Technological and organizational Perspectives, IDEA Group Inc, USA, pp. 124-143.

[7] Ávila, P., Putnik, G., Cunha, M. (2005) Broker Performance for Agile/Virtual Enterprise Integration, in Putnik, G. & Cunha, M. (Eds.) Virtual Enterprise Integration: Technological and organizational Perspectives, IDEA Group Inc, USA, pp. 166-185.

[8] Cunha, M., Putnik, G., Gunasekaran, A., Ávila, P (2005) Market of Resources as a Virtual Enterprise Integration Enabler, Putnik, G. & Cunha, M. (Eds.) Virtual Enterprise Integration: Technological and organizational Perspectives, IDEA Group Inc, USA, pp. 145-165.

[9] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," IEEE Trans Industr Inform, vol. 15, no. 4, pp. 2405–2415, 2018.

[10] A. Eyring, N. Hoyt, J. Tenny, R. Domike, and Y. Hovanski, "Analysis of a closed-loop digital twin using discrete event simulation," The International Journal of Advanced Manufacturing Technology, vol. 123, pp. 1–14, Sep. 2022, doi: 10.1007/s00170-022-10176-5.

[11] FlexSim, "Warehouse Optimization and Digital Twin | FlexSim." Accessed: Jul. 21, 2024. [Online]. Available: https://www.flexsim.com/videos/warehouse-optimization-and-digital-twin/

Biography

Afonso Silva studied industrial engineering and management at Instituto Superior de Engenharia do Porto (ISEP) and is currently pursuing a masters degree in service engineering and management at Faculdade de Engenharia da Universidade do Porto (FEUP). Currently works as a research assistant at INESC TEC at the Center for Enterprise Systems Engineering.

Paulo Ávila is an Associate Professor with Habilitation at the School of Engineering of the Polytechnic of Porto, where he previously served as Department Director. He currently directs the Bachelor's Degree in Industrial Engineering and Management and is President of FORESP – Technological School of Vale de Cambra. His research interests include resource selection, agile and virtual enterprise design, sustainability, production systems optimization, quality management, and computer-integrated manufacturing. He is an active contributor to national and international conferences, journals, edited volumes, and industry-driven projects. He is also a founding member and Co-chair of the International Conference on Business Sustainability.

João Bastos is an Associate Professor at ISEP, part of the Polytechnic Institute of Porto, and currently serves as Director of the Department of Mechanical Engineering. He holds a degree in Mechanical Engineering from the Faculty of Engineering of the University of Porto (FEUP), a Master's degree in Electrical and Computer Engineering with a specialization in Industrial Informatics from FEUP, and a PhD from the Doctoral Program in Industrial Engineering and Management, also from FEUP. His research interests include Supply Chain Management, Distributed Planning, Digital Twins, and the optimization of production systems. He is a researcher at the National Institute for Systems and Computer Engineering of Porto (INESC TEC – Associated Laboratory) and is involved in several research projects. He regularly participates in national and international conferences and publishes in scientific journals.

Luís Pinto Ferreira is an Associate Professor w/ Habilitation at Mechanical Engineering Department (Industrial Management Section) of ISEP (www.isep.ipp.pt) (Polytechnic of Porto), Porto, Portugal. He obtained his PhD in Industrial Engineering and Management area from the University of Vigo (Spain), MSc in Industrial Engineering–Logistics and Distribution from the School of Engineering, University of Minho (Portugal) and bachelor's degree in industrial electronic engineering from the School of Engineering, University of Minho (Portugal). In both the master's and doctoral studies, simulation was used as a decision-support tool in the production area. In April 2025, he obtained his Habilitation in Engineering and Industrial Management from the University of Aveiro. He presents h-index of 46 in Google Scholar (6129 citations) and 29 in SCOPUS profile (2349 citations).

Hélio Castro Hélio Castro, Dr. Sc. in Industrial and Systems Engineering, M.Sc. in Industrial Engineering and Management, Dipl. Eng. in Manufacturing Engineering at the University of Minho, Portugal. He is an Adjunct Professor, Researcher, and Sub-Director of the degree in Industrial Engineering and Management, as well as responsible for the Production Systems Laboratory at ISEP, Polytechnic of Porto, Portugal. His research interests are open-source (hardware and software) design, virtual and ubiquitous enterprises, meta-organizations, advanced manufacturing and business models, concurrent and collaborative engineering, chaos, and complexity management. He is also a senior researcher in the Enterprise System Engineering Center (CESE) at INESC TEC, Portugal, and an external collaborator in the Decision and Systemic Strategy (DECSYS) research group at the Federal University of Santa Maria, Brazil.

Alzira Mota received a PhD in Industrial and Systems Engineering from the University of Minho, Portugal, and an MSc in Artificial Intelligence and a BSc in Mathematics and Computer Science from the University of Porto, Portugal. She is an Adjunct Professor in the Department of Mathematics at the School of Engineering of the Polytechnic Institute of Porto (ISEP), Portugal, where she served as Sub-Director of the Department of Mathematics from 2012 to 2020. She is a researcher at the Enterprise Systems Engineering Centre (CESE) at INESC TEC in Portugal. Her research focuses on mathematical programming, multi-objective optimisation, evolutionary algorithms, and multi-criteria decision-making methods.