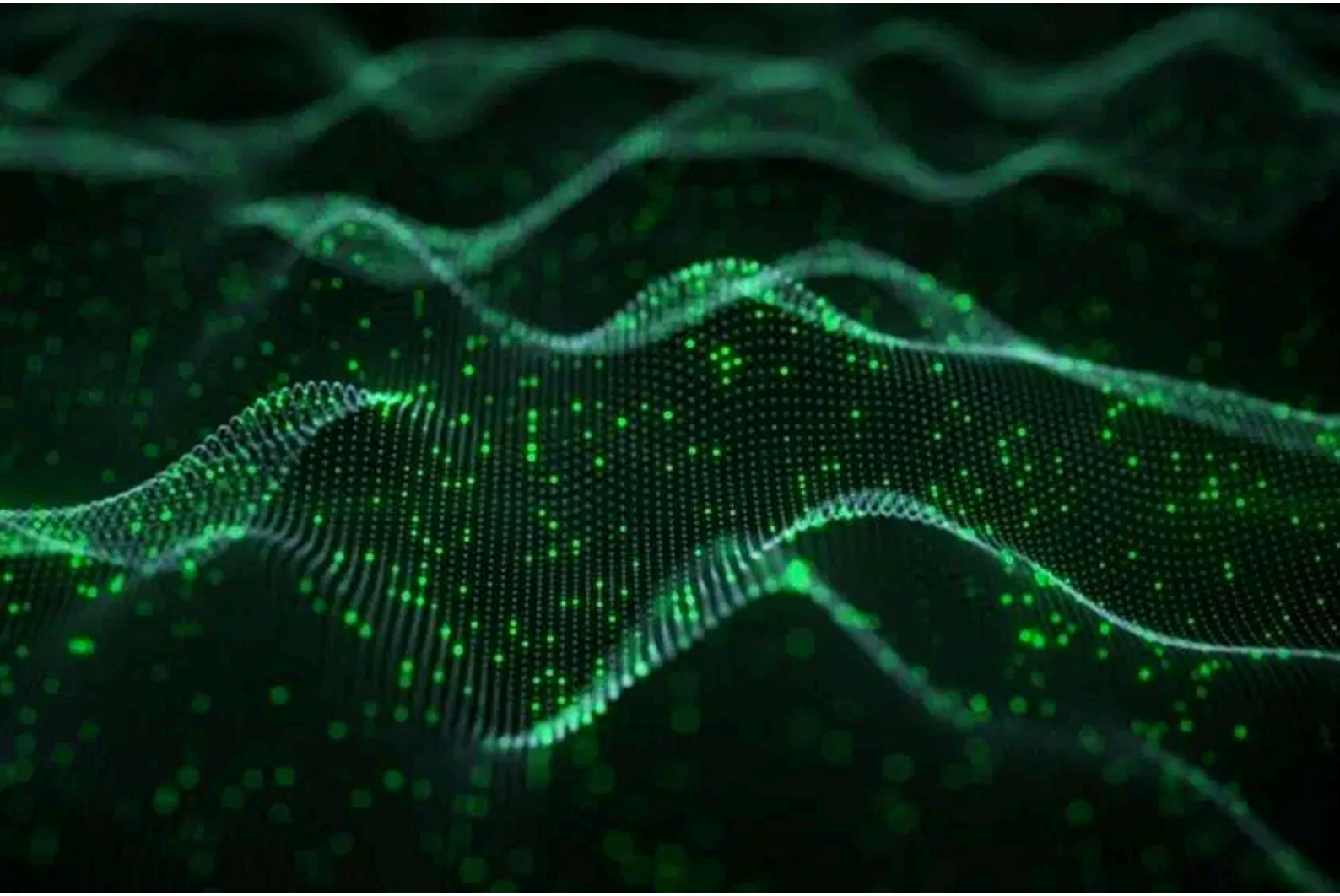
Institut für Wirtschaftsinformatik
Universität St.Gallen

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

Management of AI Lab
Institut für Wirtschaftsinformatik
Müller-Friedberg-Str. 6/8



Management von Künstlicher Intelligenz

Managing Artificial Intelligence, Workshop Paper Series, INFORMATIK 2021
(51. Jahrestagung der Gesellschaft für Informatik)

Herausgeber
Prof. Dr. Benjamin van Giffen
Prof. Dr. Jana Koehler
Prof. Dr. Walter Brenner
Prof. Dr. Can Adam Albayrak

Bitte zitieren Sie als:

van Giffen, B.; Koehler, J; Brenner, W.; Albayrak, C. (2021). Managing Artificial Intelligence, Workshop Paper Series, INFORMATIK 2021, Institute of Information Management, University of St.Gallen

*Für weitere Anfragen wenden Sie sich bitte an die entsprechenden Autoren oder an benjamin.vangiffen@unisg.ch

Management of AI Lab
Institut für Wirtschaftsinformatik
Müller-Friedberg-Str. 6/8
9000 St. Gallen, Switzerland
www.ai.iwi.unisg.ch

Table Of Contents

Vorwort

Liebe Workshopteilnehmerinnen und -teilnehmer,

Sehr geehrte Damen und Herren,

der Hype um Künstliche Intelligenz hat in den vergangenen Jahren deutlich zugenommen. Menschen in Politik und Gesellschaft, Unternehmen und öffentlichen Einrichtungen sowie auch an Universitäten und Hochschulen beschäftigen sich zunehmend mit der Frage wie KI nachhaltig im produktiven Einsatz verankert werden kann.

Die INFORMATIK 2021 steht im Zeichen der Nachhaltigkeit. Dies war ein spannender Anlass den Workshop "Management von Künstlicher Intelligenz" bei der Gesellschaft für Informatik vorzuschlagen und durchführen zu können. Als Herausgeber freuen wir uns über die positive Resonanz und vielseitigen Beiträge zum Workshop. Die bearbeiteten Themen fokussieren nicht nur auf KI-Prototypen, sondern behandeln klassische Managementfragen im Kontext der Künstlichen Intelligenz: Entwicklung von KI-Strategien, KI-Projektmanagement, Datenbeschaffung und -bereitstellung für KI-Systeme, Betrieb von KI-Infrastrukturen sowie der Entwicklung notwendiger Kompetenzen und Kultur um KI nachhaltig einzusetzen.

Natürlich kann ein so umfassendes Thema wie das Management von Künstlicher Intelligenz nicht innerhalb eines Workshops vollständig bearbeitet, geschweige denn abgeschlossen werden. Im Gegenteil, es handelt sich um ein emergentes und zugleich dynamisches Themenfeld bei dem die Zusammenarbeit unterschiedlicher Disziplinen wie Informatik und Wirtschaftsinformatik, sowie auch zwischen Wissenschaft und Praxis zentral ist.

In den Händen halten sie heute die für den Workshop ausgewählten Beiträge, welche durch Vertreterinnen und Vertreter unterschiedlicher Disziplinen und Anwendungskontexte erarbeitet wurden. Damit verbunden ist unser Anliegen den Austausch zum Thema anzuregen, um in Zukunft geeignete Methoden, Rahmenwerke und Instrumente zu entwickeln, um KI technisch, wirtschaftlich und sozial nachhaltig und erfolgreich einsetzen zu können.

An dieser Stelle möchten wir den Autoren danken für die eingereichten Beiträge und ihr Engagement im Rahmen von Präsentationen und Diskussionen im Workshop auf der INFORMATIK 2021.

Freundliche Grüsse wünschen die Herausgeber

Benjamin van Giffen

Jana Koehler

Walter Brenner

Can Adam Albayrak

Über die Herausgeber

Prof. Dr. Benjamin van Giffen

Prof. Dr. Benjamin van Giffen ist Assistenzprofessor für Wirtschaftsinformatik an der Universität St.Gallen. Am Institut für Informationsmanagement leitet das Research Lab für das Management von Künstlicher Intelligenz. Seine Forschung bearbeitet die strategischen und betriebswirtschaftlichen Fragen der Gestaltung, Skalierung und dem Betrieb von KI-gestützten Informationssystemen. Zielsetzung seiner Arbeit ist der unternehmerische, wertorientierte und verantwortungsvolle Umgang mit Künstlicher Intelligenz. Zu seinen Forschungsschwerpunkten gehören das Management von KI in Unternehmen, digitale Innovation mit KI sowie menschenzentrierte Innovation von KI-Systemen.

Prof. Dr. Jana Koehler

Sie ist seit 2019 wissenschaftliche Direktorin des Forschungsbereiches Analyse und Optimierung von Prozessen am Deutschen Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) und Inhaberin des Lehrstuhls für Künstliche Intelligenz an der Universität des Saarlandes. Sie hat Informatik und Wissenschaftstheorie an der Humboldt Universität zu Berlin studiert und 1994 an der Universität des Saarlandes in Informatik promoviert. Weitere berufliche Stationen waren die Albert-Ludwigs-Universität Freiburg, wo sie 1996 habilitierte, die Schindler AG, IBM Research Zürich und die Hochschule Luzern. Ihr Spezialgebiet sind KI-Methoden für die flexible Optimierung und Entscheidungsunterstützung in Fertigungs- und Geschäftsprozessen.

Prof. Dr. Walter Brenner

Professor Brenner joined St.Gallen University in 2001 as a professor after having held the Chair of Information Systems at the University of Essen (Germany) for two years, and at Technical University of Freiberg (Germany) for 7 years. Currently, Professor Brenner acts as Director of the Institute of Information Management. He published more than 300 Articles and more than 35 books. His research focuses on Industrialisation of Information Management, Management of IT service providers, Innovation and Technology Management and Management of Artificial Intelligence.
Professor Brenner received a graduate degree in business administration (lic. oec.) and a Doctorate (Dr. oec.) from the University of St.Gallen. Prior to joining academia Professor Brenner worked as Head of Application Development with Alusuisse-Lonza AG (Switzerland).

Prof. Dr. Can Adam Albayrak

Prof. Dr. Can Adam Albayrak arbeitete viele Jahre bei der Deutschen Post/DHL, wo er für die strategische Weiterentwicklung der IT-Bereiche zuständig war und zahlreiche IT-Projekte leitete, bevor er 2006 einem Ruf als Professor für Wirtschaftsinformatik an die Hochschule Harz folgte. Seine Lehr-, Forschungs- und Beratungsschwerpunkte liegen im operativen und strategischen IT-Management sowie im Bereich Informationsethik. Zudem leitet er den Studiengang Wirtschaftsinformatik an der Hochschule Harz, ist Autor zahlreicher Fachveröffentlichungen und ist darüber hinaus noch Lehrbeauftragter an weiteren Hochschulen. In der Gesellschaft für Informatik ist er Mitglied des Leitungsgremiums der GI-Fachgruppe IT-Controlling sowie Sprecher der GI-Fachgruppe Strategisches Informationsmanagement.

Zielsetzung des Workshops und Einführung in das Management von Künstlicher Intelligenz

Durch Management von Künstlicher Intelligenz zum nachhaltigen Erfolg von KI-Systemen

Benjamin van Giffen[1], Jana Koehler[2], Walter Brenner[1]

[1] Universität St.Gallen
[2] Universität d. Saarlandes, DFKI

Abstract: Die Algorithmen Künstlicher Intelligenz werden stetig weiterentwickelt und kommen in immer mehr Produkten und Anwendungen in Wirtschaft und Gesellschaft zum Einsatz. Zahlreiche Prototypen werden entwickelt, mit denen der Einsatz Künstlicher Intelligenz in unterschiedlichsten Anwendungsbereichen erschlossen werden soll. Dennoch gelingt nur wenigen Prototypen der Sprung in produktive Anwendungen, die einen nachhaltigen unternehmerischen Nutzen schaffen. Der Beitrag zeigt auf, dass Prozesse und Strukturen für das Management Künstlicher Intelligenz benötigt werden, um den nachhaltigen Erfolg von KI-Systemen sicher zu stellen.

Keywords: Artificial Intelligence, Information Management

1 Einleitung: Motivation und Thema

Die Algorithmen Künstlicher Intelligenz werden stetig weiterentwickelt und kommen in immer mehr Produkten und Anwendungen in Wirtschaft und Gesellschaft zum Einsatz. Zahlreiche Prototypen werden entwickelt, mit denen der Einsatz Künstlicher Intelligenz in unterschiedlichsten Anwendungsbereichen erschlossen werden soll. Dennoch gelingt weniger als 50% dieser Prototypen der Sprung in produktive Anwendungen, die einen nachhaltigen unternehmerischen Nutzen schaffen (Cearley 2021). Diese hohe Quote zeigt, dass professionelle Prozesse und Strukturen benötigt werden, um Anwendungen der Künstlichen Intelligenz zu entwickeln, zu betreiben und in den betrieblichen Kontext zu integrieren. Wir bezeichnen diese Prozesse und Strukturen als Management der Künstlichen Intelligenz.

In den verschiedensten Industriebereichen werden Projekte mit Künstlicher Intelligenz initiiert und stehen in diversen Formen auf den Agenden der Entscheiderinnen und Entscheider in Unternehmen. Die zugrundeliegenden Erwartungen und Wertversprechen sind dabei immens. Die gegenwärtigen Erfolge mit Künstlicher Intelligenz werden überwiegend durch den Einsatz Maschinellen Lernens im Rahmen von Proof-of-Concepts, in ersten Anwendungen und in Produkten wie Sprachassistenten realisiert. Durch diese Machbarkeitsnachweise im realen Anwendungskontext der Unternehmen werden Erfahrungen im Umgang mit der Technologie gemacht sowie deren Besonderheiten erlernt, beziehungsweise teilweise schmerzlich erfahren.

Da derzeit viele der Projekte, in denen Künstliche Intelligenz zum Einsatz kommt, scheitern, kann von einer flächendeckenden Implementierung nicht gesprochen werden. Eine Grundvoraussetzung für die Implementierung Künstlicher Intelligenz in industriellen Anwendungsbereichen ist der verlässliche und stabile Betrieb. Eine Klarheit hinsichtlich der Regelung von Verantwortung und Haftung bei fehlerhaften Systementscheidungen und daraus resultierenden Implikationen muss von Anfang an gegeben sein. Das Management von Künstlicher Intelligenz ist eine Erweiterung des bestehenden Managements der Informatik.

Die INFORMATIK 2021 steht im Zeichen der Nachhaltigkeit. Die nachhaltige Entwicklung bzw. der nachhaltige Einsatz von Künstlicher Intelligenz stellt Wissenschaftlicher und Praktiker vor zahlreiche Herausforderungen. Die Neuartigkeit von KI-Systemen erfordert eben auch einen anderen Umgang als mit traditionell programmierter Software. Voraussetzung für den nachhaltigen Einsatz von Künstliche Intelligenz ist, dass KI-Systeme systematisch entwickelt und betrieben werden, anstatt dem starken Trend “Prototypenbau” hinterherzulaufen. KI-Anwendungen sollen schließlich ökonomisch, sozial, technologisch und auch ökologisch tragfähig sein. Die Frage ist jedoch, wie eine integrierte Sicht aussehen kann und soll, und weiterhin auch, wie der Weg dorthin durch fundierte Managementmethoden geebnet werden kann.

Zielsetzung des Workshops

Zielsetzung des Workshops ist es, relevante Aspekte eines nachhaltigen Managements von Künstlicher Intelligenz zu beleuchten. Dieses betrifft u.a. KI-Strategieentwicklung, KI-Projektmanagement, aber auch Herausforderungen in Bezug auf unternehmensweite Kompetenzen, den Betrieb geeigneter technischer Infrastrukturen, sowie die Entwicklung notwendiger Fähigkeiten und der Unternehmenskultur für die Systementwicklung, als auch den Betrieb von KI-Lösungen (vgl. Abbildung 1).

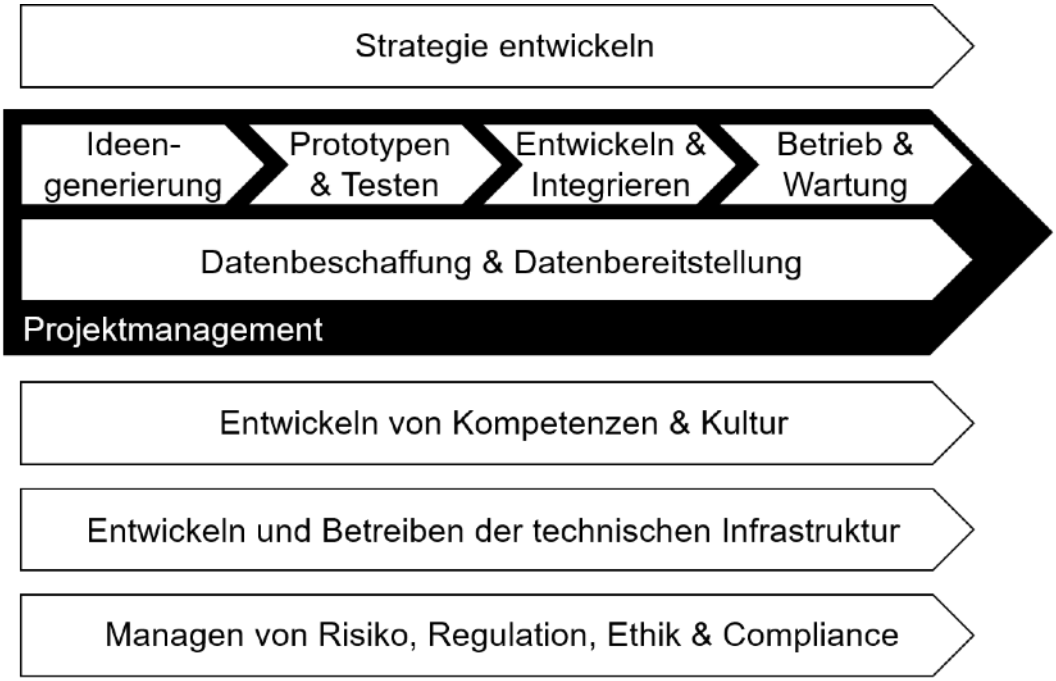


Abbildung 1: Prozesse des Managements Künstlicher Intelligenz

Gegenstand des Workshops sind spezifische Herausforderungen in den o.g. Themenbereichen. Hierbei sollen zum einen Lösungsansätze und Best Practices vorgestellt und diskutiert werden, zum anderen ein Forum für den interdisziplinären Austausch zwischen Informatikern und Wirtschaftsinformatikern als auch Theoretikern und Praktikern geschaffen werden.

Für den Workshop wurden insgesamt sieben Forschungsarbeiten, darunter insbesondere Fallstudien, konzeptionell orientierte Papiere, sowie Research in Progress Artikel ausgewählt, so dass ein interaktiven Austausch im Workshop möglich wird. Die Beiträge decken unterschiedlichen Aspekte des oben gezeigten, rudimentären Prozessmodells für das Managements Künstlicher Intelligenz ab.

Datenbeschaffung und Datenbereitstellung

Beitrag 1 “Managerial Cultivation of Training Data for Artificial Intelligence Systems on the Consumer Internet of Things” von Jan Zibuschka beschäftigt sich Beschaffung und Bereitstellung von Daten für das Training von KI-Modellen. Der Autor greift auf zwei KI-Implementierungsprojekte zurück, um Herausforderungen für die Datenverfügbarkeit zu identifizieren.

Management von Risiko, Regulation und Compliance

Beitrag 2: “Developing IT Systems Based on Machine Learning: Specific Risks and Mitigation Strategies” von Benedikt Berger, Johann Kranz, Ioanna Constantiou und Thomas Hess entwirft ein Forschungsdesign für die Untersuchung von KI-spezifischen Risiken im Rahmen der Informationssystementwicklung.

KI-Projektmanagement

Beitrag 3: "Wie die AUDI AG datenbasierte Technologien für die Serienproduktion entwickelt und erprobt: Das Audi Production Lab" von André Sagodi, Henning Löser und Verena Bossdorf zeigt auf wie die AUDI AG mit dem P-Lab, einer dedizierten Einheit für das Innovations- und Technologiemanagement in der Automobilfertigung, systematisch KI-Innovationen in den produktiven Betrieb überführt.

Entwickeln von Kompetenzen und Kultur

Beitrag 4: "Künstliche Intelligenz für die Smart City – Handlungsimpulse für die kommunale Praxis" von Tabea Hein und Christian Schachtner zeigen Wege auf, wie KI-gestützte digitale Innovationen in der öffentlichen Verwaltung durch die Entwicklung von KI-Kompetenzen und Kultur gefördert werden können.

Management der technischen Infrastruktur

Beitrag 5: "AI Lab – An Ecosystem for Managing AI projects" von Martin Leucker und Maria Ostanina beschreiben mit dem AI Lab ein KI-Ökosystem, welches Werkzeuge und Ressourcen für die Entwicklung von KI-Systemen bündelt und bereitstellt. Ein Schwerpunkt der Arbeit liegt in der nahtlosen Unterstützung des Machine Learning Workflows, d.h., insbesondere für das Trainieren und den Einsatz von Machine Learning Modellen, sowie das zugehörige Prozessieren von Trainingsdaten.

Beitrag 6: "Towards Reference Architectures for Artificial Intelligence: A Preliminary Report" von Alexis T. Bernhard, Jana Koehler und Sophia Saller beleuchtet mögliche Referenzarchitekturen um gängige KI-Anwendungsszenarien zu beschreiben. Vor diesem Hintergrund bestehende Architekturstile, die für KI-Systeme vorgeschlagen wurden (z. B. die Lambda-Architektur, Blackboard-Architektur und Agenten-Architekturen), vorgestellt und im Hinblick auf Reifegrad und mögliche Weiterentwicklungen reflektiert.

Beitrag 7: AI-ware: Bridging AI and Software Engineering for responsible and sustainable intelligent artefacts" von Dagmar Monett und Claudia Lemke weist auf die Notwendigkeit eines differenzierten Verständnis von "Künstliche Intelligenz" hin, d.h. insbesondere diesen nicht nur auf Verfahren des Maschinellen Lernens zu beschränken. Darüber hinaus unterstreichen die Autorinnen die zunehmend relevante Verbindung von KI- und Softwareentwicklung in betrieblichen Umgebungen hin.

Ausblick: Management von Künstlicher Intelligenz

Es handelt sich um den ersten Workshop zum Management von KI im Rahmen der INFORMATIK-Konferenz. Der Workshop gliedert sich in eine Reihe von Veranstaltungen im Bereich der Betriebswirtschaftslehre bzw. Wirtschaftsinformatik ein: Der Verband der Hochschullehrer für Betriebswirtschaftslehre (VHB) hatte im November 2020 eine Arbeitstagung "Beyond the Hype: Künstliche Intelligenz in Wissenschaft und Praxis" organisiert. Auf der International Conference of Information Systems (ICIS) in Dezember 2020 gab es einen Practice Development Workshop (PDW) mit dem Titel "AI Beyond the Hype". Der im Rahmen der INFORMATIK 2021 durchgeführte Workshop knüpft an diese Veranstaltungen an, soll die Thematik aber insbesondere in Richtung der Informatik öffnen und ein Forum für Informatikerinnen und Informatiker bieten, die an Management-Fragestellungen von KI interessiert sind und sich hierzu aktiv austauschen möchten. Mittel- bis langfristig ist es angedacht, Folge-Workshops anzubieten und diese auf konkrete Problemstellungen zu fokussieren.

3 Referenzen und Vorarbeiten zum Thema

Albayrak, C. A., Renn, O., und Teille, K. 2018. „Leitlinien für das menschliche Handeln in einer digitalisierten Welt“, HMD Praxis der Wirtschaftsinformatik (55:5), Springer Fachmedien Wiesbaden GmbH, S. 1048–1064.

Barenkamp, M., Rebstadt, J., und Thomas, O. 2020. „Applications of AI in classical software engineering“, AI Perspectives (2:1), aiperspectives.springeropen.com, S. 1.

Berente, N., Gu, B., Recker, J., und Santhanam, R. 2019. „Managing AI“, Call for Papers, MIS Quarterly, misq.org. (https://www.misq.org/skin/frontend/default/misq/pdf/CurrentCalls/ManagingAI.pdf).

Berger, B., Adam, M., Rühr, A., und Benlian, A. 2020. „Watch Me Improve—Algorithm Aversion and Demonstrating the Ability to Learn“, Business & Information Systems Engineering, Springer. (https://doi.org/10.1007/s12599-020-00678-5).

Brenner W., van Giffen B., Koehler J. (2021) Management of Artificial Intelligence: Feasibility, Desirability and Viability. In: Aier S., Rohner P., Schelp J. (eds) Engineering the Transformation of the Enterprise. Springer, Cham. https://doi.org/10.1007/978-3-030-84655-8_2

Brenner, W., van Giffen, B., Koehler, J., Fahse, T., und Sagodi, A. 2021. Bausteine eines Managements Künstlicher Intelligenz: Eine Standortbestimmung, Essentials, Springer.

Buxmann, P., Hess, T., und Thatcher, J. B. 2020. „AI-Based Information Systems“, Business & Information Systems Engineering. (https://doi.org/10.1007/s12599-020-00675-8).

Buxmann, P., und Schmidt, H. 2018. Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg, Springer-Verlag.

Cearley, D. 2021. The Gartner 2021 Top Strategic Technology Trends. (https://www.gartner.com/en/webinars/3993288/the-gartner-2021-top-strategic-technology-trends?ref=btem&utm_campaign=RM_GB_YOY_WBNR_BB_E1_48HR&utm_medium=email&utm_source=Eloqua&cm_mmc=Eloqua-_-Email-_-LM_RM_GB_YOY_WBNR_BB_E1_48HR-_-0000).

van Giffen, B., Borth, D., und Brenner, W. 2020. „Management von Künstlicher Intelligenz in Unternehmen“, HMD Praxis der Wirtschaftsinformatik, Springer. (https://link.springer.com/article/10.1365/s40702-020-00584-0).

Hofmann, P., Jöhnk, J., Protschky, D., und Urbach, N. 2020. „Developing purposeful AI use cases--a structured method and its application in project management“, in 15th International Conference on Wirtschaftsinformatik (WI). (https://www.fim-rc.de/Paperbibliothek/Veroeffentlicht/1025/wi-1025.pdf).

Koehler, J. 2018. „Business Process Innovation with Artificial Intelligence: Levering Benefits and Controlling Operational Risks“, European Business & Management (4:2), S. 55.

Koehler, J., van Giffen, B., und Brenner, W. 2021. „Zieldimensionen für ein erfolgreiches KI-Projektmanagement“, Handelsblatt, S. 14–16.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., und Söllner, M. 2019. „AI-based digital assistants“, Business & Information Systems Engineering (61:4), Springer, S. 535–544.

Monett, D., Lewis, C. W. P., Thórisson, K. R., Bach, J., Baldassarre, G., Granato, G., Berkeley, I. S. N., Chollet, F., Crosby, M., Shevlin, H., Fox, J., Laird, J. E., Legg, S., Lindes, P., Mikolov, T., Rapaport, W. J., Rojas, R., Rosa, M., Stone, P., Sutton, R. S., Yampolskiy, R. V., Wang, P., Schank, R., Sloman, A., und

Winfield, A. 2020. „Special Issue “On Defining Artificial Intelligence”—Commentaries and Author’s Response“, Journal of Artificial General Intelligence (11:2), Berlin: Sciendo, S. 1–100.

Beiträge zum Workshop

Datenbeschaffung und Bereitstellung

Managerial Cultivation of Training Data for Artificial Intelligence Systems on the Consumer Internet of Things

Jan Zibuschka[1]

[1] Robert Bosch, GmbH, Zentralbereich Forschung und Vorausentwicklung, Renningen, 70465 Stuttgart, jan.zibuschka@de.bosch.com

Abstract: Systems based on the use of Artificial Intelligence (AI) are on the rise and impact many aspects of everyday life. One important use case of such systems is controlling the myriad of networked devices commonly referred to as the Internet of Things (IoT). While IoT devices produce voluminous data that trigger the use of AI, availability of suitable information to use as training data during development is a common and significant impediment to the success of projects in this area. Based on experience in several industry research projects focusing on consumer IoT use cases, we describe the relevance of this training information, reasons that lead to an inadequate availability, and steps taken to alleviate the issue.

Keywords: Artificial Intelligence, Information Management, Internet of Things

1 Introduction

The Internet of Things (IoT) has arrived in consumer products. Leading manufacturers in application areas from household appliances to automotive are integrating network interfaces into their products. The reasons for this are manyfold. The reduced cost of networking components makes it economically viable to include them as a differentiator for high-end products in a wide range of use cases, while in other cases regulators require networked systems for their promise of increasing safety. Processing the high volume, unstructured information produced by such deployments commonly requires the use of Artificial Intelligence (AI) [Zi19b]. Those AI-based systems on the IoT are commonly referred to as assistance systems, and the convergence of the two technology streams has been summarized with a combined acronym: the AIoT [Ka19].

Underlying the AI in these systems is machine learning (ML), which has seen significant advances in recent years [Am19]. These novel machine learning technologies power modern assistance systems in products such as voice assistants, but are also the main component in AI software platforms [Am19], which are central to business success on the AIoT [Ka19]. Correspondingly, novel engineering challenges in such systems are related to machine learning, with availability of training information named as the most significant challenge by a wide array of practitioners [Am19]. While this scarcity of training data certainly has a significant impact on engineering activities, in this contribution we consider it as a managerial issue. Lack of training information is an impediment to the success of AI development, which is often organized in agile teams [Lw19]. As AI systems typically require large amounts of training information to parametrize them before the perform adequately [Gi20], tackling this issue should have a high priority for managers.

2 Case Studies

This contribution describes experiences with the availability of training information for AI-based IoT systems targeted at consumers from several industry projects. In addition to internal industry projects we cannot discuss in great detail here, which also limits the findings of related work describing findings at Microsoft [Am19], the challenges and mitigation approaches draw on experience in two public, consortial research projects that have been amply publicized.

- The ENTOURAGE project [Zi19a] investigated an ecosystem of digital assistants in several IoT use cases, including smart home and smart mobility. ENTOURAGE brought together an automotive industry association, a major industry player covering several IoT application domains, and several smaller enterprises. The smart mobility area was mainly focused on connected cars, multimedia, and driver assistance, however, there was also some work on autonomous driving systems.
- The ForeSight project [Ba19] aims to develop an AI platform for smart living scenarios. The consortium includes industry associations for electrical engineering and housing, and a large number of industry partners of various sizes as either consortium members or associated partners. The smart living domain goes beyond the smart home into application domains such as energy management, however, the authors were mainly involved in the smart home space.

As is evident by the references we provided for the discussion in the previous sections, our findings from our case studies broadly align with experiences from other settings [Am19, Lw19].

3 Challenges for the Availability of Training Information

Impediments to the availability of training information may stem from many sources. In our case studies, the main challenges we found could be categorized as follows:

- Compliance and Ethical Issues: Issues of compliance and ethics are well known to impact the development of AI systems, and also significantly impact the availability of training information in consumer IoT use cases [SW18]. Specifically, issues of privacy and data protection may affect the availability of training information in AIoT scenarios, as sensor data in many cases, at least potentially, contains personal information of varying and unclear criticality. This may cause compliance issues, ethical issues, and may be a material concern for consumers' purchasing decisions [Zi19b]. Conversely, in some cases training information may be needed to assess ethical and compliance issues, such as the degree to which personal information is involved or necessary, or the availability of information enabling the training of non-discriminating machine learning models.
- Unprecedented Use Cases: Due to the highly innovative nature of AI use cases present in many consumer IoT use cases, the information needed to train the ML components at the core of those use cases are simply not available, and may never have been collected as corresponding sensor arrays may not have been deployed in corresponding settings. Without any data, the AI system will perform very poorly, an engineering problem referred to as cold-start [Pl19]. On a behavioral level, this may lead to changed behavior of the first users of a system, leading to significant skew in the learned models [Li01]. Finally, on a market level, this may lead to a chicken-egg problem, as consumers do not purchase systems with poor training and performance, and therefore those systems cannot be adequately trained.
- Organizational Issues: Another significant hurdle regularly encountered in the availability of training data for AI systems are organizational boundaries. It has long been recognized that sharing of information across organizational boundaries, even within one organization requires dedicated, encouraging management structures [Wa05]. As a result, even when training data is available in an organization, it may not be available where it is needed [Am19]. Such issues do however not necessarily result from inadequate management, and could also be an expression of transaction costs and resource organization as underlying drivers of organizational structure [SE05], hinting that it might not be sensible to force the issue, as this might go against an efficient organizational structure in some cases

4 Managerial Cultivation of Training Information

Three main stakeholder groups determine the success of an AI project: end users, engineers, and managers [Ko21]. In this contribution, we focus on the managerial contribution. Unlike engineering approaches, which lead to immediate results upon implementation, the managerial approach calls for a more long-term cultivation mindset, as is common in knowledge and information management issues [LR07]. Approaches to cultivation of training information for artificial intelligence we observed in our case studies include:

- Non-critical deployments: Managers should push for deployment of early versions of AI systems in non-critical scenarios, which may offer opportunities to collect training information while avoiding many of the pitfalls associated with deployment of undertrained systems. Examples for this include living labs [Al12] or corresponding test vehicles for mobility use cases. Such settings

enable an incremental user involvement, and thus enable the collection of training data under relatively realistic conditions with a desired degree of intervention to address possible bias issues, while also allowing for evaluation of highly innovative systems.

- Role and Competence Management: Having appropriate competences in appropriate roles within the projects also proved instrumental in facilitating availability of training information. One example for this are data protection compliance issues with acquiring training information. To enable the AI team to communicate with the data protection hierarchy, knowledge in the areas of privacy and data protection is instrumental. However, data protection officers are obliged to promote privacy in an organization, which may lead to a perspective that perceives AI as a problem per se. Having independent competence in an AI team may lead to a development of privacy-aware AI that alleviates these concerns, and will enable a deeper understanding of data protection concerns by the AI team. Another similar example is knowledge management across organizational boundaries, where it is key to have competences in integration of heterogenous information and a managerial understanding of standardization problems in information management.
- Consortial Approach: Finally, embedding the AI activity in a consortial context, especially with involvement of industry associations, has the potential of allowing access to training information for AI-based systems via partners in the consortium. While such inter-organizational collaboration needs to be approached carefully from a compliance point of view, it may even be more efficient than collaboration within a structured organization as described above. The acquisition of data from various settings also has the potential to reduce bias and improve accuracy of the ML models.

5 Discussion

While the cultivation strategies we propose certainly come with their own set of challenges, motivating further research, we identified the potential of management approaches for tackling the availability of training information beyond the purely technological approaches described by earlier work [Am19]. We feel our experience could aid practitioners in addressing similar challenges, and academics to identify research directions with the potential for a major impact in industry. There is also significant potential for extending or research. In this short contribution, we could only give a coarse overview of the impediments and approaches we observed most frequently, without in-depth analysis. We would also like to see comparisons of the approaches we saw in consumer-focused application domains to other domains such as AI in industry or energy management: the issue of availability of training data should generalize to those application domains, while the managerial approaches we propose might not.

Bibliography

[Al12] Almirall, E. et al.: Mapping Living Labs in the Landscape of Innovation Methodologies. Technology Innovation Management Review. September 2012: Living Labs, 12–18 (2012).

[Am19] Amershi, S. et al.: Software Engineering for Machine Learning: A Case Study. In: 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). pp. 291–300 (2019).

[Ba19] Bauer, J. et al.: ForeSight - Platform Approach for Enabling AI-based Services for Smart Living. In: Pagán, J. et al. (eds.) How AI Impacts Urban Living and Public Health. pp. 204–211 Springer International Publishing, Cham (2019).

[Gi20] van Giffen, B. et al.: Management von Künstlicher Intelligenz in Unternehmen. HMD. 57, 1, 4–20 (2020). https://doi.org/10.1365/s40702-020-00584-0.

[Ka19] Katz, J.S.: AIoT: Thoughts on Artificial Intelligence and the Internet of Things, https://iot.ieee.org/conferences-events/wf-iot-2014-videos/56-newsletter/july2019.html 2019.html, last accessed 2021/03/15.

[Ko21] Koehler, J. et al.: Zieldimensionen für ein erfolgreiches KI-Projektmanagement. Handelsblatt. (2021).

[Li01] Liebowitz, J.: Knowledge management and its link to artificial intelligence. Expert Systems with Applications. 20, 1, 1–6 (2001).

[LR07] Loyarte, E., Rivera, O.: Communities of practice: a model for their cultivation. Journal of Knowledge Management. 11, 3, 67–77 (2007).

[Lw19] Lwakatare, L.E. et al.: A taxonomy of software engineering challenges for machine learning systems: An empirical investigation. In: International Conference on Agile Software Development. pp. 227–243 Springer, Cham (2019).

[Pl19] Pliakos, K. et al.: Integrating machine learning into item response theory for addressing the cold start problem in adaptive learning systems. Computers & Education. 137, 91–103 (2019). https://doi.org/10.1016/j.compedu.2019.04.009.

[SE05] Santos, F.M., Eisenhardt, K.M.: Organizational boundaries and theories of organization. Organization science. 16, 5, 491–508 (2005).

[SW18] Stahl, B.C., Wright, D.: Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation. IEEE Security Privacy. 16, 3, 26–33 (2018). https://doi.org/10.1109/MSP.2018.2701164.

[Wa05] Walczak, S.: Organizational knowledge management structure. The Learning Organization. 12, 4, 330–339 (2005). https://doi.org/10.1108/09696470510599118.

[Zi19a] Zibuschka, J. et al.: The ENTOURAGE Privacy and Security Reference Architecture for Internet of Things Ecosystems. In: Open Identity Summit 2019. Gesellschaft für Informatik, Bonn (2019).

[Zi19b] Zibuschka, J. et al.: Users' Preferences Concerning Privacy Properties of Assistant Systems on the Internet of Things. In: AMCIS 2019 Proceedings. (2019).

Author information

Dr. Jan Zibuschka

Dr.-Ing. Jan Zibuschka is a Research Engineer at Robert Bosch GmbH, in the central research department's division for advanced digital technologies. His research interests include security and privacy, the internet of things, data science, machine learning, and information systems in general. In his work, he focuses on human and strategic factors in the use of technology. He has authored dozens of scientific publications and holds over thirty patents.

Management von Risiko, Regulation und Compliance

Developing IT Systems Based on Machine Learning: Specific Risks and Mitigation Strategies

Benedikt Berger[1], Johann Kranz[2], Ioanna Constantiou[2], Thomas Hess[2]

[1]Department of Information Systems, University of Münster

[2] LMU Munich School of Management, Ludwig-Maximilians-Universität München

Abstract: Machine learning has become increasingly powerful in solving problems in a wide array of applications areas. Organizations seek to utilize this power by implementing machine learning in their information technology systems to enhance their business. However, the development of such systems is subject to specific challenges, which entail potential risks to the systems' successful implementation and use. While the existing literature provides rich insights on the risk management in information technology development, we know little about the specific risks caused by the employment of machine learning. In this research-in-progress paper, we lay the foundations for a study that seeks to fill this gap.

Keywords: Information technology development, machine learning, risk management

[2] Department of Digitalization, Copenhagen Business School

1 Introduction

The use of machine learning (ML) has led to substantial performance improvements of various information technology (IT) applications [An17]. In contrast to explicitly writing software code, ML allows to realize IT capabilities by training statistical models based on existing data or feedback gathered over time. The increasing amounts of data and power of processors required for training as well as advancements of the models to be trained have fuelled this development. Accordingly, ML is supposed to spur further technological progress.

Despite the ML-related achievements and promises, the development of IT systems employing ML is prone to failure [EEV21]. Prominent examples of such failure include a recruiting tool developed by Amazon that discriminated female applicants and a chatbot developed by Microsoft that propagated racist content. While failure of IT projects is a common phenomenon, the reasons for failure of IT projects involving ML likely differ from those that do not involve ML owing to differences in the development process [Zh20]. The functionality of ML-based systems depends largely on the quality of the training data and the configuration of the underlying models usually involves some extent of experimentation. It stands to reason that the probabilistic nature of ML, in contrast to the deterministic nature of software engineering, introduces additional uncertainties and risks to the successful realization of ML-based systems.

Although software engineering, information systems (IS) development and risk management in IT projects are subject of a large body of literature [HM18; MR19], our knowledge about the ML development process and the management of uncertainties and risks throughout this process is scarce. Therefore, this research-in-progress paper outlines a study that seeks to answer the following research question: Which ML-related risks occur along the development of ML-based systems and how can companies manage these risks?

To answer this research question, we propose a research design employing a case study approach with IT projects involving ML as cases. Before explaining this design in detail, we introduce the conceptual foundations of our analysis, which comprise the ML development process and the literature on risk management in IT projects.

2 Conceptual Foundations

2.1 Machine Learning Development Process

The development process of ML-based systems differs from the widely established planbased or agile software development practices [Wa19]. Most existing representations describe this process as a partly iterative workflow, pipeline, or lifecycle, which we summarize in five phases (see Fig. 1): problem definition, data preparation, model development, deployment, and maintenance [Am19; Hi16; Hu19; LK18; Qi20]. During the problem definition, the developers build an understanding of the problem they are supposed to solve, define the requirements of the system to be build, and narrow done the type of ML model applicable to the problem. The data preparation entails the collection, cleaning, and, if necessary, labelling of the data. Depending on the problem, the data may stem from multiple internal or external sources and usually requires transformation to enable integration and ingestion into the model. To begin the model development, developers must decide, which features (i.e. properties contained in the data) the statistical model should draw upon to solve the problem. Moreover, the developers must define the chosen model's hyper-parameters (e.g. the number of hidden layers in an artificial neural network) before training the model based on the data and evaluating the trained model's performance. Because the feature selection and hyperparameter configuration have a strong influence on the performance, these process steps are usually subject to iterations to experimentally optimize these model characteristics. Once the evaluation yields satisfactory results, the model can be deployed, which means it processes new data (i.e. data that

was not part of the model training). After deployment, the model's performance should be continuously monitored to identify the need for retraining, which we name maintenance.

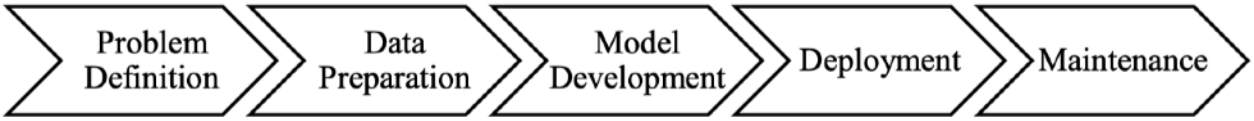


Fig. 1: ML Development Process

The differences in the development of IT systems employing ML also lead to specific challenges in the development process [Wa19]. These begin with the problem definition, which must account for the differences in how ML and rule-based systems solve problems [Wa19; Zh20]. Moreover, the management of large amounts of complex and volatile data differs substantially from the management of code, which is the focus of established software development methods [Am19] and corresponds to most developers skills [Wa19]. The context-specificity of the training procedure also exacerbates the reuse of ML-based systems in different contexts [Am19; EEV21] as well as the separate development and subsequent integration of modules [Am19]. Furthermore, the experimental character of the model development and the probabilistic nature of ML models introduce inherent risks to the development process [EEV21]. The opacity of many ML models additionally hampers their evaluation because their output may lack interpretability and, thus, justification in case of doubts [Zh20]. Finally, the performance of ML-based systems may degrade over time, which requires continuous and, if possible, automated maintenance [Wa19]. Overall, these challenges call for a rigorous investigation of the potential risks in ML development processes and how they can be managed.

2.2 Risk Management in IT Development Projects

In the context of IT development projects, risks constitute "threats to success" [Sc01]. In other words, risks are potential determinants of project failure. The exposure of a project to a specific risk is the product of the likelihood that the threat materializes and the impact this event would have on the project [Bo91]. Following this conceptualization, the risk management process comprises the assessment of risks, including their identification, analysis, and prioritization, and the control of risks, including response planning, response execution, and monitoring [Bo91; TAW12].

Early research on risk management in IT development projects sought to identify and cluster risks as well as their sources [e.g. Mo96]. Major risk dimensions include the requirements of the IT system to be developed, the technologies that are necessary to realize the system, the project team and management, the users and further stakeholders of the system, the organization (e.g. the corporate culture and support of the project), and the market environment [Li10; Sc01; Ta06; WKR04]. While the risks of using ML in an IT development project can be understood as a technology risk, we propose that this understanding does not reflect the substantial changes in the development process required by ML. Moreover, we are not aware of studies that consider data to be a risk factor. However, the fundamental role of data in the ML process calls for such investigations.

The risk management literature further comprises two rather separated research streams [MR19]: the normative body of IT project risk management, which seeks to optimize risk management decisions, and the experiential body, which investigates how IT project managers actually behave. The normative literature prescribes methods, such as scorecards or checklists, to identify, assess, and mitigate IT projects risks [e.g. BRT93]. For the purpose of risk mitigation, the literature suggests appropriate risk responses depending on the risk source [LMR98]. If the risk responses fit the risk sources, they are supposed to reduce risk exposure and increase project performance [BRT01].

However, the experiential literature on IT project risk management has shown that IT project manager often disengage from formal risk management practices [KH05]. This is because their risk perception may lead to different conclusions than formal risk analysis. This tension between prescribed and enacted IT project risk management comes down to the question whether intuition

or deliberate analysis is superior in dealing with a specific risk, which likely depends the risk's and the context's characteristics [MR19]. Despite the extensive existing considerations of these matters, the literature provides insights about neither how ML-related risks should be dealt with nor whether managers of IT projects involving ML actually perceive such risks. Our study seeks to undertake first steps in this direction.

3 Study Design

To answer our research question, we plan to conduct a multiple case study with IT projects involving ML as the unit of analysis. Our sampling strategy aims at creating a heterogeneous set of projects, differing in the following aspects:

- Age of the implementing company (start-ups vs. established firms)
- Digital origin of the implementing company (digital-born vs. transforming)
- Parties involved in the development process (in-house vs. external)
- Importance of ML in the system to be implemented (marginal vs. core feature)
- Life cycle of the development process (short-term vs. long-term)

We think that such a broad sampling approach allows us to investigate the development of ML-based systems in the required breadth to abstract from the specific circumstances of specific cases. For data collection, we plan to draw upon interviews with project participants from all parties involved in the project, on-sight observations as well as project documentation.

Bibliography

[Am19] Amershi, S. et al.: Software Engineering for Machine Learning: A Case Study. In: Proceedings of the 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP), Montreal, Quebec, Canada 2019, pp. 291-300.

[An17] Anthes, G.: Artificial Intelligence Poised to Ride a New Wave. Communications of the ACM 60/7, pp. 19-21, 2017.

[Bo91] Boehm, B. W.: Software risk management: principles and practices. IEEE Software 8/1, pp. 32-41, 1991.

[BRT93] Barki, H.; Rivard, S.; Talbot, J.: Toward an Assessment of Software Development Risk. Journal of Management Information Systems 10/2, pp. 203-225, 1993.

[BRT01] Barki, H.; Rivard, S.; Talbot, J.: An Integrative Contingency Model of Software Project Risk Management. Journal of Management Information Systems 17/4, pp. 3769, 2001.

[EEV21] Engel, C.; Ebel, P.; van Giffen, B.: Empirically Exploring the Cause-Effect Relationships of AI Characteristics, Project Management Challenges, and Organizational Change. In: Proceedings of the 16th International Conference on Wirtschaftsinformatik (WI), Essen, Germany 2021.

[Hi16] Hill, C. et al..: Trials and tribulations of developers of intelligent systems: A field study. In: Proceedings of the 2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), Cambridge, United Kingdom 2016, 162-170.

[Hu19] Hummer, W. et al.: ModelOps: Cloud-Based Lifecycle Management for Reliable and Trusted AI. In: Proceedings of the 2019 IEEE International Conference on Cloud Engineering (IC2E), Prague, Czech Republic 2019, pp. 113-120.

[HM18] Hassan, N. R.; Mathiassen, L.: Distilling a body of knowledge for information systems development. Information Systems Journal 28/1, pp. 175-226, 2018.

[KH05] Kutsch, E.; Hall, M.: Intervening conditions on the management of project risk: Dealing with uncertainty in information technology projects. International Journal of Project Management 23/8, pp. 591-599, 2005.

[Li10] Liu, S. et al.: Comparing senior executive and project manager perceptions of IT project risk: a Chinese Delphi study. Information Systems Journal 20/4, pp. 319-355, 2010.

[LK18] Lui, K.,; Karmiol, J.: AI Infrastructure Reference Architecture. IBM Systems, Research Triangle Park, North Carolina, United States, 2018.

[LMR98] Lyytinen, K.; Mathiassen, L.; Ropponen, J.: Attention Shaping and Software Risk—A Categorical Analysis of Four Classical Risk Management Approaches. Information Systems Research 9/3, pp. 233-255, 1998.

[Mo96] Moynihan, T.: An inventory of personal constructs for information systems project risk researchers. Journal of Information Technology 11/4, pp. 359-371, 1996.

[MR19] Moeini, M.; Rivard, S.: Sublating Tensions in the IT Project Risk Management Literature: A Model of the Relative Performance of Intuition and Deliberate Analysis for Risk Assessment. Journal of the Association for Information Systems 20/3, pp. 243-284, 2019.

[Qi20] Qian, B. et al.: Orchestrating the Development Lifecycle of Machine Learning-based IoT Applications: A Taxonomy and Survey. ACM Computing Surveys 53/4, Article 82, 2020.

[Sc01] Schmidt, R. e al.: Identifying Software Project Risks: An International Delphi Study. Journal of Management Information Systems 17/4, pp. 5-36, 2001.

[Ta06] Taylor, H.: Risk Management and Problem Resolution Strategies for it Projects: Prescription and Practice. Project Management Journal 37/5, pp. 49-63, 2006.

[TAW12] Taylor, H.; Artman, E.; Woelfer, J. P.: Information Technology Project Risk Management: Bridging the Gap between Research and Practice. Journal of Information Technology 27/1, pp. 17-34, 2012.

[Wa19] Wan, Z. et al.: How does Machine Learning Change Software Development Practices? IEEE Transactions on Software Engineering, in press, 2019.

[WKR04] Wallace, L.; Keil, M.; Rai, A.: Understanding software project risk: a cluster analysis. Information & Management 42/1, pp. 115-125, 2004.

Author information

Dr. Benedikt Berger

Benedikt Berger is an assistant professor of digital transformation and society at the Department of Information Systems at the University of Münster, Germany. His current research focuses on digital media and AI-based information systems. His work has appeared in the Journal of Management Information Systems, Electronic Markets, Business & Information Systems Engineering and in various international conference proceedings.

Prof. Ioanna Constantiou

Ioanna Constantiou is a professor of information systems at the Department of Digitalization in Copenhagen Business School, with special responsibility to conduct research on Digital Transformation. Her current research work focuses on the impact of digital platforms in traditional industries as well as the impact of digital technologies and in particular AI in everyday life and the digital transformations in society. Her research has been published in a number of academic outlets, including, the Journal of Information

Prof. Dr. Thomas Hess

Thomas Hess is director of the Institute for Information Systems and New Media at LMU Munich School of Management. His research focuses on the intersection between management and digitization, the digitization of media companies, and AI-based information systems. His work has been published in the Journal of Management Information Systems, the European Journal of Information Systems and Business & Information Systems Engineering, among others

Prof. Dr. Johann Kranz

Johann Kranz is a professor of digital services and sustainability at the LMU Munich School of Management. His research focuses on IS-business alignment and governance in the digital age, and green IS for enabling circular economies, sustainable mobility and pro-environmental behavior. His research has been published in the Journal of Strategic Information Systems, the Information Systems Journal, MIS Quarterly Executive, the Journal of Service Research and Energy Policy.

KI-Projektmanagement

Wie die AUDI AG datenbasierte Technologien für die Serienproduktion entwickelt und erprobt: Das Audi Production Lab

André Sagodi[1,2], Henning Löser[1], Verena Bossdorf[1]

[1] AUDI AG, Production Lab, Ingolstadt, Deutschland
[2] Universität St.Gallen, IWI-HSG, St. Gallen, Schweiz, andre.sagodi@student.unisg.ch

Abstract: Datenbasierte Technologien wie Künstliche Intelligenz bieten immenses Wertpotenzial für produzierende Unternehmen. Für die Bewertung dieser Potenziale im spezifischen Unternehmenskontext und die Erarbeitung von Lösungen für den produktiven Betrieb bedarf es eines dedizierten Innovations- und Technologie-managements. An dieser Stelle ist bei der AUDI AG das Production Lab tätig. Dieser Beitrag skizziert die Rolle des Audi Production Labs und thematisiert vielversprechende Anwendungsbereiche, aktuelle Herausforderungen und relevante Fragestellungen rund um den Einsatz von Künstlicher Intelligenz in der Produktion.

Keywords: Datenbasierte Technologien; Künstliche Intelligenz; Industrie 4.0

1 Technologischer Wandel im industriellen Umfeld

Im Verlauf des letzten Jahrhunderts hat die fertigende Industrie mehrere industrielle Revolutionen durchlebt, in denen neue technologische Innovationen zu tiefgreifenden Veränderungen in den Fabriken geführt haben [Dw21]. Mechanisierung, Massenproduktion und Automatisierung haben das Fertigungsumfeld geprägt und Prozesse wurden in den letzten Jahrzehnten unter Berücksichtigung von Prinzipien des Lean Manufacturing, Six Sigma und Total Quality Management standardisiert, modularisiert und im erforderlichen Maße flexibilisiert. In diesem Produktionsumfeld sind die Aufgaben für Mensch und Maschine klar definiert, Spezifikationen werden überwacht und Leistungskennzahlen wie Stückzahl, Zykluszeit, Stillstandzeit oder die Gesamtanlageneffektivität (OEE) werden zur Kontrolle und Steuerung herangezogen. Eine zentrale Herausforderung in der heutigen Automobilproduktion ist der Umgang mit der zunehmenden Individualisierung und Produktvarianz. Kunden können Fahrzeuge nach Wunsch konfigurieren und beispielsweise zwischen einer Vielzahl komplexer, optionaler Features wie autonomen Fahrfunktionen wählen. Dies erfordert den Einbau zusätzlicher Hardware- und Softwarekomponenten in die Fahrzeuge, was in den Produktionsprozessen berücksichtigt werden muss. Um die damit immer komplexer werdenden Produktionslinien wirtschaftlich zu betreiben sind flexible Automatisierungslösungen sowie die Digitalisierung von Produkten und Prozessen entscheidende Wettbewerbsfaktoren [VD18]. Daher investieren Unternehmen heute stark in datengetriebene Technologien wie Big Data, Cloud und Künstliche Intelligenz. Im deutschsprachigen Raum werden diese primär IT-getriebene Veränderungen in den Fertigungssystemen als Industrie 4.0 zusammengefasst [La14]. Die Entwicklung und Erprobung dieser neuen Technologien für den Einsatz im hoch performanten Produktionsumfeld ist eine entscheidende Fähigkeit und ein wichtiger Wettbewerbsfaktor für Unternehmen.

2 Datenbasierte Technologien in der Produktion

Datenbasierte Technologien bieten vielfältige Wertpotenziale für produzierende Unternehmen. Die Fortschritte insbesondere im Bereich von Künstlicher Intelligenz (KI) werden das Tempo der Automatisierung weiter beschleunigen [BM17] und es wird erwartet, dass dieser weltweite Trend branchenübergreifend große geschäftliche Auswirkungen auf Unternehmen haben wird [GA19]. So wird der Automatisierungsgrad in der verarbeitenden Industrie schätzungsweise von weltweit circa 10 Prozent auf etwa 25 Prozent ansteigen [BPJ21]. In Kombination mit der weiter steigenden Datenverfügbarkeit wird sich dies wahrscheinlich grundlegender auf Unternehmen und deren Belegschaft auswirken als die industrielle Revolution [KH19]. Der Begriff ‚Künstliche Intelligenz' ist ein teils irreführender Oberbegriff, der sich vor allem auf Anwendungen aus dem Bereich des maschinellen Lernens (Machine Learning) bezieht [LW21] und bereits in den 1950er Jahren geprägt wurde. KI-Technologien ermöglichen die Automatisierung von Funktionen, die bisher nur mit kognitiven menschlichen Fähigkeiten ausgeführt werden konnten, wie beispielsweise das Erfassen, Wahrnehmen und Interagieren mit der Umwelt, Problemlösen, Lernen oder Entscheiden [BPJ21, RSC19]. Einsatzszenarien in der Fertigung sind beispielsweise die prädiktive Wartung, die Vorhersage von Störungen, das automatisierte Entscheiden in der Produktionssteuerung, Materialdisposition, Qualitätsüberwachung oder auch die (teil-) automatisierte Prozessoptimierung. Nachfolgend werden exemplarisch drei zentrale Anwendungsbereiche datenbasierter Technologien in der Produktion erläutert:

- **Qualitätsüberwachung:** Hier haben KI-Algorithmen ein großes Potential. Das Einsatzspektrum umfasst die optische Prüfung hinsichtlich der verbauten Teilevariante oder die Qualitätskontrolle, wie beispielsweise die automatische Detektion von Rissen in Tiefziehteilen im Presswerk [AU18]. Weitere Anwendungsbeispiele sind die Vorhersage der Güte von Schraubverbindungen oder das Detektieren von Rast-Geräuschen bei Steckverbindungen.
- **Intelligente Planung komplexer Systeme:** Durch die erhebliche Steigerung der Komplexität in der Produktion zusammen mit dem Bedarf kürzerer Prognose- und Reaktionszeiten ist eine

umfassende Planung unter Einbezug aller Einflussgrößen manuell durch den Menschen oder mit klassischer Planungs-Logik kaum mehr abbildbar. Hier kann KI einen Mehrwert bieten, indem verschiedene Planungsalternativen unter Berücksichtigung komplexer Eingangsgrößen aufgezeigt und bewertet werden. Dies soll zu einer geglätteten Auslastung, erweiterten Berücksichtigung verschiedener Eingangs- und Einflussgrößen und damit letztlich zu Prozessen mit gesteigerter Stabilität, Sicherheit und Effizienz führen.

- **Entscheidungsunterstützung:** Aufgrund der Komplexität der Zusammenhänge in einer verketteten Fertigung ist die Fehlerpropagation durch den Fertigungs- und Logistikprozess für einen Menschen nicht beherrschbar. Hierfür können neue Methoden entwickelt werden, die den regulären Ablauf einer Produktion und den Logistikprozess im Hinblick auf Prozessreihenfolge, Output und zu erwartende Qualität interpretieren. Dabei sind Faktoren, wie Auslastung, Varianz und der Faktor Mensch zu berücksichtigen, um auftretende Schwankungen in der lose verketteten Produktion zu identifizieren und deren Einfluss auf das Gesamtsystem abzuschätzen. Die datengestützte Anomalie-Erkennung und der automatisierte Abgleich mit ähnlichen Vorfällen kann die Schwere von Störungen im Fertigungsablauf erheblich verringern.

3 Das Audi Production Lab

Das Audi Production Lab (P-Lab) wurde 2012 im Ingolstädter Stammwerk der AUDI AG gegründet und bietet Infrastruktur und organisatorisch verankerte Start-upGedankenkultur für Innovationen für die weltweite Audi-Produktion. Dies ermöglicht die Erprobung neuer Technologien und die Erarbeitung von produktionsrelevanten Use Cases. Das P-Lab forciert dabei die Identifikation von sowohl geschäfts- als auch technologiegetriebenen Anwendungsfällen und ordnet diese in den größeren strategischen Kontext des Unternehmens ein. Außerdem ermöglicht das P-Lab das Management von externen Innovationen und stellt die Vernetzung zu externen Anbietern, Start-ups und Universitäten her. Die heutige Zahl der circa 30 festen Mitarbeiterinnen und Mitarbeiter multipliziert sich durch temporäre Nutzerinnen und Nutzer, Projektbeteiligte und Promotoren um ein Vielfaches [AU21].

Abb. 1: Das Audi P-Lab als Reißverschluss zwischen Idee und Serienfertigung

Die Entwicklung und Implementierung neuer technologischer Innovationen bei Audi dient nicht dem Selbstzweck, sondern muss letztendlich zu einer Wertrealisierung im produktiven Betrieb führen [AU21]. Das bedingt den Übergang von explorativen Aktivitäten zu verwertenden Geschäftsprozessen. Unter Berücksichtigung begrenzter finanzieller und personeller Ressourcen beinhaltet dies auch eine Priorisierung von Technologien und die Ressourcenzuweisung zu ausgewählten, besonders relevanten Projekten. Umso wichtiger ist es Experten aus den Fachbereichen, also potenzielle Nutzer, bereits frühzeitig in die Aktivitäten zu integrieren, um deren Domänenwissen und fachliche Expertise zu berücksichtigen. In einer frühen Phase ist das Ziel des P-Labs zunächst kein fertiges Produkt, sondern der Aufbau von Fähigkeiten und Wissen über geeignete Technologien, die Durchführung von Vorüberlegungen und die frühzeitige Validierung von potenziellen Anwendungsfällen. Dies ermöglicht die Bewertung von Use Cases mit einem höheren Reifegrad sowie eine genauere Kostenschätzung im Zuge der Gestaltung eines ganzheitlichen

Technologie-Portfolios. Nach dem Motto „fail early, fail often, fail forward" ist das Scheitern eines Versuchs durchaus eingeplant; ebenso wie die kurzfristige Inbetriebnahme eines neuen Testaufbaus. Dieses explorative Vorgehen wäre in einer produktiven Umgebung nicht möglich, da es der Logik eines stabilen, optimierten und fehlerfreien Betriebs widerspricht [AU21].

3.1 Innovationen für die Produktion

Die Aktivitäten in der frühen Innovationsphase können sowohl problemgetrieben (Problem sucht Technologie/Lösung) als auch technologiegetrieben (Technologie sucht Problem) sein. Die resultierenden Use Cases unterscheiden sich in Zielsetzung und Reifegrad vom „Entwickeln und Testen" über die „Qualitätsverfeinerung und Validierung" bis hin zum „produktiven Einsatz und Betrieb". Deshalb muss das Ziel eines Use Cases von Anfang an klar definiert werden – geht es um den Beweis, dass etwas technologisch möglich ist oder darum, dass bestimmte Kriterien wie beispielsweise Genauigkeit, Wiederholrate, Verfügbarkeit oder Kostengrenze erreicht werden können? Die Erfahrung zeigt, dass gute Problem-Lösungs-Paarungen nicht „einfach so" entstehen und zudem häufig verschiedene Technologien grundsätzlich für die Lösung in Frage kommen. Deren Tauglichkeit für den produktiven Einsatz unter gegebenen kontextuellen Bedingungen muss dann jeweils erst erprobt und nachgewiesen werden. So können beispielsweise Probleme in der Produktion oft auf verschiedene Arten angegangen werden: (1) einige können und sollten mit IT gelöst werden, (2) einige können, sollten aber nicht mit IT gelöst werden, (3) einige sind organisatorischer Natur und nicht mit Technologie lösbar und (4) einige müssen erst tiefgreifender verstanden werden, bevor das weitere Vorgehen definiert werden kann. Zudem sind es nicht unbedingt die naheliegenden, offensichtlichen Lösungen, die am Ende zum Erfolg führen. Für diesen kreativen, explorativen Prozess ist die Berücksichtigung von Fach- und Technologiewissen sowie eine technologieoffene und kritische Herangehensweise essenziell, die nur im Umfeld einer offenen Innovationskultur entstehen und gelebt werden kann - dies bietet das Audi P-Lab [AU16].

3.2 Fragestellungen beim Einsatz Künstlicher Intelligenz

Das breite Einsatzspektrum datenbasierter Technologien bietet vielfältige Möglichkeiten im Bereich der Produktion zu innovieren. Zudem fallen bereits heute zahlreiche Daten an, deren Informationen die Grundlage für innovative Projekte bilden können. Gleichzeitig kann der breite Einsatz dieser Technologien nur in die Produktion integriert werden, wenn die technologischen Eigenschaften und Anforderungen ganzheitlich in der Organisation verstanden werden. Der produktive Einsatz von angewandter KI in Unternehmen und insbesondere in der industriellen Produktion geht weit über die reine Forschung hinsichtlich Algorithmen, Methoden und Toolkits hinaus und ist mit vielen weiteren Aspekten und Fragestellungen verbunden. Das betrifft neben dem Aufbau von Know-how und Kompetenzen vor allem auch Themen wie industrietaugliche Infrastrukturen, Plattformen, Standardisierung, Bereitstellung beziehungsweise Zugänglichkeit von Daten, Sicherstellung hinreichender Datenqualität, maschinelle Lesbarkeit der Daten, Steigerung der Reife von KI-Anwendungen, Skalierbarkeit der Lösungen sowie Datenschutz und Informationssicherheit. Nachfolgend werden zentrale Handlungsfelder und Fragestellungen beim Einsatz von KI in der Produktion erläutert:

- Für die automatisierte Qualitätsüberwachung werden je nach Use Case unterschiedliche Daten als Eingangsgrößen herangezogen, die in der Regel durch Anwendung trainierter klassischer Algorithmen der Mustererkennung die gewünschten Ergebnisse liefern. Für einen breiten Einsatz in der Produktion müssen zusätzlich einfache Methoden zum Trainieren von Algorithmen zur Mustererkennung entwickelt werden. Dabei muss sichergestellt werden, dass bei der Integration von neuen Fahrzeugvarianten oder Änderungen im Prozess, z.B. durch KVP-Maßnahmen, das Nachtrainieren der Algorithmen ohne Einbezug (externer) Spezialisten realisiert werden kann. Neu ist hierbei für eine klassische Fertigung, dass KI-Algorithmen auf

Basis von Wahrscheinlichkeitswerten zu Entscheidungen führen, was neben technischen auch rechtlichen Fragen aufwirft: *Ist die Verbau-Dokumentation auf Basis von Mustererkennung und Wahrscheinlichkeiten zulässig?*

- Im Gegensatz zur Qualitätsüberwachung liegen bei der Prozessoptimierung beziehungsweise der Prozessüberwachung selten genügend Daten von Störungen vor, um einen Algorithmus auf diese zu trainieren. Hierfür sind neue Ansätze zu entwickeln, die über Metriken zur Bewertung von Anomalien verfügen: *Ist dieser Ausreißer in den Prozessdaten relevant oder nur eine Anomalie in den Messdaten? Gab es diesen Ausreißer schon einmal? Wird sich die Anomalie über den weiteren Prozess (in der Produktion) fortpflanzen?* Auf Basis dieser Metriken könnten dann entweder Mitarbeiter auf die Anomalie aufmerksam gemacht werden, den Mitarbeitern Handlungsempfehlungen unterbreitet werden oder es kann sogar automatisiert gegengesteuert werden.

- Eine weitere unbeantwortete Fragestellung ergibt sich aus der Betrachtung der Frage nach dem „deployment" und dem „Life Cycle Management" der Algorithmen. In einer klassischen Speicherprogrammierbaren Steuerung (SPS) werden circa 1.000 Variablen alle 10 ms aktualisiert. In einer Fertigungslinie gibt es circa 1.000 SPS. Das ergibt zusammen eine Million Variablen, die mit 100 Hz ausgelesen und verarbeitet werden müssen. Für diese Anwendungsfälle müssen neue Verfahren entwickelt werden, um die „feature extraction" und „data ingestion" nahe am Entstehungsort der Rohdaten (zum Beispiel direkt in der SPS) durchzuführen, damit anstelle von großen Mengen an Rohdaten nur die relevanten Informationen übertragen werden. Diese Verfahren zur Informationsextraktion müssen dabei flexibel genug sein, um mit wachsender Prozesskenntnis einfach angepasst und in die Fertigungsstätten verteilt werden zu können, ohne die Fertigung zu beeinflussen.

- Der Einsatz von KI-Systemen zielt in der Fertigung auf Effizienzsteigerungen ab. *Wie funktioniert aber das Life Cycle Management oder die Detektion von Fehlverhalten in den KI-Systemen und wie werden Fehler und Störungen in diesen Systemen behoben?* Hierzu müssen die Mitarbeitenden selbst befähigt werden und für den Betrieb und die Instandhaltung von KI-Systemen geschult werden. Das umfasst das Verständnis der notwendigen Technik zur Datenübertragung, Datenmodellierung der Prozesse und das Verständnis der möglichen Fehler. Die Eigendiagnose von KISystemen ist heute größtenteils unbekannt, aber dennoch ein wichtiger Faktor für den produktiven Einsatz. Es müssen neue Konzepte zur geführten Fehlersuche gefunden werden, die die Eigendiagnose der KI-Systeme nutzt.

- Aufgrund der teilweise radikalen Änderungen in den innerbetrieblichen Abläufen ist ein Aus- und Weiterbildungskonzept für den Aufbau, die Einrichtung, die Benutzung, Wartung und Instandhaltung von KI-Systemen essenziell. Neben dem handwerklichen Geschick der Mitarbeiter wird zukünftig ein Wissen über die Fehlerbehebung von automatisierten Prozessen noch wichtiger werden. Hierzu müssen die Inhalte der Ausbildungsberufe betrachtet und wo nötig ergänzt werden.

- Klassische Fertigungen arbeiten mit Verfügbarkeiten größer 97,5 Prozent; eine Reduzierung dieser Verfügbarkeit durch komplexere Wartung oder schwierigere Fehleranalyse und Fehlfunktionsbehebung würde die gewünschten Effizienzsteigerungen durch den Einsatz von KI schnell zunichtemachen. Einzelprozesse haben durch die Weiterentwicklungen der letzten Jahrzehnte Prozessqualitäten erreicht, in denen Fehler so selten sind, dass man Jahre bräuchte, um genügend Daten von Fehlern zu sammeln, mit denen ein Algorithmus trainiert werden kann. Deshalb ist es neben der Optimierung der Mustererkennungsmethoden in einem Produktions- und Logistikkontext notwendig die Verfahren des unüberwachten Lernens (unsupervised learning) auf diesen Kontext zu adaptieren und Metriken zur Bewertung von Anomalien in multivariaten Umgebungen zu entwickeln. Neben hoher Expertise in den Verfahren der KI setzt dies ein fundiertes Wissen über eine klassische Produktion und Logistik voraus sowie einen anwendungs- und wertorientierten Ansatz.

4 Zusammenfassung und Ausblick

Der technologische Fortschritt per se ist kein Garant für die Erreichung von Geschäftszielen [GHP14] und Technologieinvestitionen per se sind kein verlässlicher Indikator für das Maß technologiegetriebener Veränderungen in einem Unternehmen [Or00]. Der Mehrwert für die Wertschöpfung in einem Unternehmen entsteht vielmehr durch dessen Fähigkeit moderne (Automatisierungs-) Technologien zu verbinden [LW21]. Die Fähigkeit eines Unternehmens einen guten ‚fit' zwischen Technologien und Geschäftsprozessen in einem gegebenen organisatorischen Kontext herzustellen kann dabei über Erfolg und Misserfolg entscheiden [BW20]. Auch der Fortschritt datengetriebener Technologien manifestiert sich nicht nur durch Innovationen in den Technologien selbst, sondern insbesondere durch Innovationen bei der Anwendung und Implementierung dieser Technologien in produktive Geschäftsprozesse. Wenn Organisationen vom Einsatz von KI profitieren wollen, benötigen sie ein differenziertes Verständnis der Werkzeuge und der damit verbundenen Herausforderungen; zum Beispiel müssen sie vermeiden die potenziellen Kosten von KI zu unterschätzen [CC20]. In dem Zusammenhang ist es entscheidend sich mit datengetriebenen Technologien wie KI auseinanderzusetzen und deren Wertpotenzial im Kontext, mit Bezug auf spezifische Aufgaben, zu untersuchen und im Vergleich zu alternativen Lösungen zu bewerten. Mit dem Production Lab hat die AUDI AG eine separate Umgebung geschaffen, in der die Erkundung und Ausarbeitung neuer Technologien und Lösungen für den produktiven Betrieb nach Maßgabe der Audi-Anforderungen vorangetrieben wird, ohne dabei geschäftskritisch für produktive Wertschöpfungsprozesse zu sein – ein wichtiger Baustein auf dem Weg zur datengetriebenen Produktion.

Literaturverzeichnis

[AU16] AUDI AG, https://www.audi-mediacenter.com/de/audi-techday-smart-factory7076/das-audi-production-lab-7080, Stand: 17.11.2016

[AU18] AUDI AG, https://www.audi-mediacenter.com/de/pressemitteilungen/audi-optimiertqualitaetspruefung-im-presswerk-mit-kuenstlicher-intelligenz-10847, Stand: 15.10.2018

[AU21] AUDI AG, https://www.progress.audi/progress/de/audi-production-lab.html, Stand: 29.01.2021

[BW20] Bednar, P. M.; Welch, C.: Socio-Technical Perspectives on Smart Working: Creating Meaningful and Sustainable Systems. Information Systems Frontiers, 22(2), S. 281-298, 2020.

[BPJ21] Benbya, H.; Pachidi, S.; Jarvenpaa, S. L.: Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. Journal of the Association for Information Systems, 22(2), S. 281-303, 2021.

[BM17] Brynjolfsson, E.; Mitchell, T.: What can machine learning do? Workforce implications. Science, 358(6370), S. 1530-1534, 2017.

[CC20] Canhoto, A. I.; Clear, F. Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. Business Horizons, 63(2), S. 183-193, 2020.

[Dw21] Dwivedi, Y. K., et.al.: Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 2021.

[GHP14] George, G.; Haas, M. R.; Pentland, A.: Big Data and Management. Academy of Management Journal, 57(2), S. 321-326, 2014.

[Ga19] Gartner, https://blogs.gartner.com/smarterwithgartner/top-trends-on-the-gartner-hypecycle-for-artificial-intelligence-2019/, Stand: 12.09.2019 Kaplan, A.; Haenlein, M.: Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons, 62(1), S. 15-25, 2019.

[KH19] Kaplan, A.; Haenlein, M.: Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons, 62(1), S. 15-25, 2019.

[LW21] Lacity, M.; Willcocks, L.: Becoming Strategic with Intelligent Automation. MIS Quarterly Executive, 20(2). https://doi.org/10.17705/2msqe.000XX, 2021.

[La14] Lasi, H., et.al.: Industry 4.0. In Business & Information Systems Engineering (Vol. 6, Issue 4, S. 239-242). https://doi.org/10.1007/s12599-014-0334-4, 2014.

[Or00] Orlikowski, W. J. (2000). Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. Organization Science, 11(4), 404-428.

[VD18] VDMA, http://normung.vdma.org/documents/22594015/26907148/Entwurf%20VDMA%2066412-40_2018-12_1539172669625.pdf/70551746-a825-b716-84b8-9540eb622e6c, 2018.

Author information

André Sagodi

André Sagodi has studied business and engineering at the Erlangen-Nuremberg University and currently works on his PhD on the topic of managing Artificial Intelligence for use in industrial manufacturing as a part-time doctoral student at the University of St.Gallen. He joined AUDI AG in 2014 as an international trainee focusing on production and plant planning. Subsequently, he was responsible for projects in the area of assembly planning. Since December 2019, André is working at the Audi Production Lab as a technology and innovation manager in the areas of assembly and logistics.

Dr. Henning Löser

Dr. Henning Loeser has studied physics as a fellow of the “Studienstiftung des deutschen Volkes” at the Philipps University Marburg and the UIUC in Urbana Champaign in Illinois, USA, and made his PhD on the topic of high temperature diffusion on single crystal surfaces at the University of Marburg. In 2004 he joined AUDI AG, starting in the area of car development superstructure and later changed to production development. He is a member of the senior management since 2012. Since November 2014 Henning is working on Audi’s future vision of the Smart Factory and is the head of the Audi Production Lab as of January 2016.

Verena Bossdorf

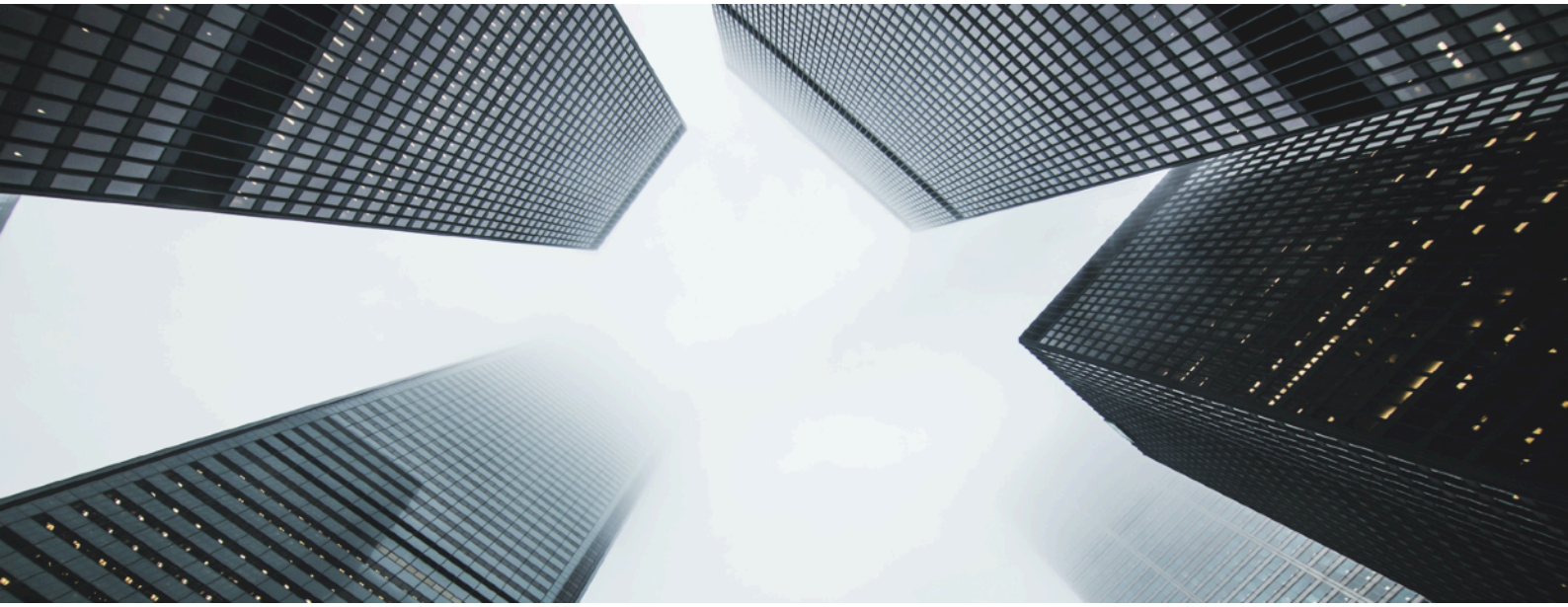
Verena Bossdorf has studied international economics at Maastricht University in the Netherlands and the University of Zaragoza. In 2003 she joined the Volkswagen Group in the purchasing department in Wolfsburg, and after a four-year assignment in Beijing, she started at AUDI AG in 2008. Since 2012 she is a member of the management. After more than a decade in procurement, Verena joined the Audi Production Lab in 2018. She and her team are responsible for technology management in the areas of assembly and logistics, ranging from scouting of new technologies to the implementation of use cases in the real production environment.

Entwickeln von Kompetenzen und Kultur

Künstliche Intelligenz für die Smart City
Handlungsimpulse für die kommunale Praxis

Tabea Hein[1] und Christian Schachtner[2]

[1] Stadt Frankfurt am Main, Koordinierungsstelle für Aufgaben der Verwaltungsstrukturreform, Berliner Straße 33-35, 60311 Frankfurt, tabea.hein@stadt-frankfurt.de
[2] IU Internationale Hochschule · Fernstudium, Studiengangleiter Public Management, Kaiserplatz 1, 83435 Bad Reichenhall, christian.schachtner@iu.org



Abstract: Künstliche Intelligenz (KI) kann ein wichtiger Lösungsbaustein zu medienbruchfreien, digitalen Services der Smart City der Zukunft sein. Während KI aus dem Alltag längst nicht mehr wegzudenken ist und der Transformationsdruck auf die öffentliche Verwaltung weiter steigt, findet KI in der Kommunalverwaltung allerdings bisher nur in geringem Maß Verwendung. Digitale Verwaltungsinnovationen können zielführend realisiert werden, wenn Stadtverwaltungen sich den Herausforderungen einer unbeständigen und komplexen Welt stellen, Wissen und Handlungskompetenzen erwerben, Synergien vielfältig nutzen und die anstehenden Aufgaben zum Wohl aller mit einer Portion Mut angehen.

Keywords: Künstliche Intelligenz, Smart City, Stadt, Kommunalverwaltung, Öffentliche Verwaltung, Basiswissen, Grundlagen, Handlungsimpulse, Praxis, Best Practices

1 Digitale Verwaltungstransformation

Die digitale Transformation der Gesellschaft schreitet unter der Prämisse „schneller, komplexer, vernetzt und disruptiv“ weiter voran. Durch Kommunikationswege, veränderte Arbeits- und Produktionsprozesse wie neue Geschäftsmodelle und Märkte sind KI-Komponenten in IT-Lösungen längst selbstverständlicher Teil unseres Alltags geworden. Adäquat privater Dienstleister erwartet die Stadtgesellschaft eine moderne, zukunftsgerichtete und bürgernah agierende Verwaltung. Forderungen nach digitalen Informationskanälen, schnellem Kommunikationsaustausch, medienbruchfreier Bearbeitung kundenfokussierten Dienstleistungen, die einfach, zeit- und ortsunabhängig nutzbar sind, sowie gemeinwohlorientierter Partizipation und Teilhabe sind dabei wesentliche Bestandteile.

Als Front- und Backend des föderalen Systems steht die Kommunalverwaltung als Garantin von Rechtstreue und Verlässlichkeit vor der Herausforderung in einer unbeständigen Welt mit schwer vorhersehbaren Ereignissen auf vielschichtigen Entscheidungsebenen in ihrer ambigen Rolle gleichzeitig Zuverlässigkeit auszustrahlen und auf gesellschaftliche wie technische Veränderungen adäquat zu reagieren. Die Coronakrise, die fortdauernde Arbeit an den OZG-Leistungen, potenzielle Auswirkungen des Registermodernisierungsgesetzes, Prozessoptimierungen zur Digitalisierung, Smart City-Projekte, Open Data und Nachhaltigkeit sind nur einige der Schlagworte, die den aktuellen, herausfordernden Aufgabenmix kommunaler Verwaltung umschreiben. Sie erzeugen Handlungsdruck, dem sich die öffentliche Hand nicht entziehen kann. Der Einsatz von KI kann als ein Lösungsbaustein zur Verwaltungsmodernisierung neue Wege eröffnen.

2 Chancen und Potenziale

KI-Lösungen bieten als Bestandteil von Digitalisierung zahlreiche Chancen und vielfältige Einsatzmöglichkeiten über nahezu die gesamte Leistungspalette der Kommunalverwaltungen. Sie können vorteilhaft in Verwaltungsarbeit integriert werden und gegenüber klassischer IT einen entscheidenden Wirkungsunterschied bedeuten. Chancen und Potenziale bieten sich vielfältig. Mitarbeitende können von seriellen Aufgaben entlastet und freiwerdende Zeitslots zu Beratungstätigkeit oder Projektmanagement eingesetzt werden. Medienbrüche können überwunden, Prozesse effizienter gestaltet werden, Partizipation und Teilhabe gefördert werden. Assistenzsysteme können dem Fachkräftemangel entgegenwirken, verminderte Prozesskosten können zu einer geringeren Abgabenlast führen. Weitere positive Auswirkungen für die Stadtgesellschaft können u.a. die Reduzierung von Behördengängen, die Beschleunigung von Antragsverfahren und die Förderung von Barrierefreiheit und Inklusion sein. Viele KI-Lösungen sind schon durch andere Branchen erarbeitet und können teils mit wenigen Anpassungen auf die kommunale Verwaltung übertragen werden.

Annäherungen an KI finden in Kommunen partiell als Bestandteil von Digitalisierungsinitiativen statt (z. B. Chatbots), praktische Umsetzungen sind jedoch durch verschiedene Faktoren wie u.a. rechtliche Rahmenbedingungen, gesellschaftliche Aushandlungsprozesse zu Ethik, datenschutzrechtliche Bedenken, mangelndes Wissen und Kompetenzen, limitiert. Gerade Letzteres steht der Annäherung an das Themengebiet KI entgegen oder kann zu Bedenken, Ängsten und Widerständen führen.

Der Einsatz von KI-Lösungen bedingt neues Wissen und Fachkompetenzen über das Themengebiet „Digitalisierung“ hinaus. Da KI-Lösungen keine „Fertigprodukte“ sind, ist ein nicht unerheblicher Anteil an Eigenleistung in Projekt und Betrieb zu erbringen. Schulungen für Führungskräfte und Mitarbeitende sind daher notwendig um mittelfristig den Aufbau eigener Expertise sicherzustellen. Datenanalytische Expertise ist wichtig, um ein grundsätzliches Verständnis von KI zu erlangen, Ausschreibungen vorbereiten zu können, mit Dienstleistern zu interagieren und als Ansprechpartner für die Stadtgesellschaft auskunftsfähig zu sein. Auch im Betrieb sollten fachverantwortliche Personen die Qualität der KI-Lösung einschätzen und zumindest teilweise Anpassungen

eigenständig vornehmen können - ohne in jedem Fall dazu Externe zu beauftragen. KI-Projekte sind im Grunde erst dann in behördliche Abläufe integriert und gelungen.

3 Kompetenzaufbau

Um Projekte erfolgreich zu realisieren, sollten sich die beteiligten Führungskräfte und Mitarbeiter:innen einer Kommunalverwaltung vorab mit den wichtigsten Aspekte und Besonderheiten des Einsatzes von KI befassen. Dieser komplexe Lernvorgang kann mithilfe des E-Books „Künstliche Intelligenz für die Smart City - Handlungsimpulse für die kommunale Praxis“ [HV20] unter der Creative Commons-Lizenz „CC BY-NC-ND 3.0 DE“ stark vereinfacht werden. Ein besonderes Anliegen der Publikation ist, branchenspezifisch und adressatengerecht einen einfachen Einstieg in die Thematik zu ermöglichen, Basiswissen zu vermitteln und einen praxisnahen Beitrag zu leisten. Sprachlich verfolgt sie dabei den Balanceakt zwischen fachlicher Präzision und Allgemeinverständlichkeit. Ziel ist es, den Leser zu befähigen, die Einführung und Anwendung von KI in der Stadtverwaltung anzustoßen und erste eigene Schritte zu gehen.

Beginnend mit Grundlagen der künstlichen Intelligenz werden Begriffe und Abgrenzungen erläutert sowie die essentielle Rolle der Daten hervorgehoben. Nutzen und Potenziale von KI werden perspektivisch zu Herausforderungen wie Ethik, Datenschutz und Vergaberecht in Beziehung gesetzt. Als ein initiales Projekt zum Einstieg in das Thema KI und zur Beschreibung von Technologie und Projektansatz dient exemplarisch ein Chatbot. Neben vielen praktischen Hinweisen wird der Beschreibung möglicher Use Cases für verschiedenste kommunale Services über die gesamte Bandbreite der Smart City-Themenfelder viel Raum gegeben. Der Fokus liegt dabei ausgehend von technischen Ansätzen auf konkreten Anwendungsbereichen sowie realistischen Umsetzungsideen, auch im Hinblick auf Übertragbarkeit, Skalierbarkeit und künftige Entwicklungen. Zum Nutzen der Kommunalverwaltung und der Stadtgesellschaft finden sich vielfältige Anwendungen der Bild- und Messdatenverarbeitung, der Sprach- und Textverarbeitung, in der Analyse komplexer Daten sowie in Vorhersage und Planung.

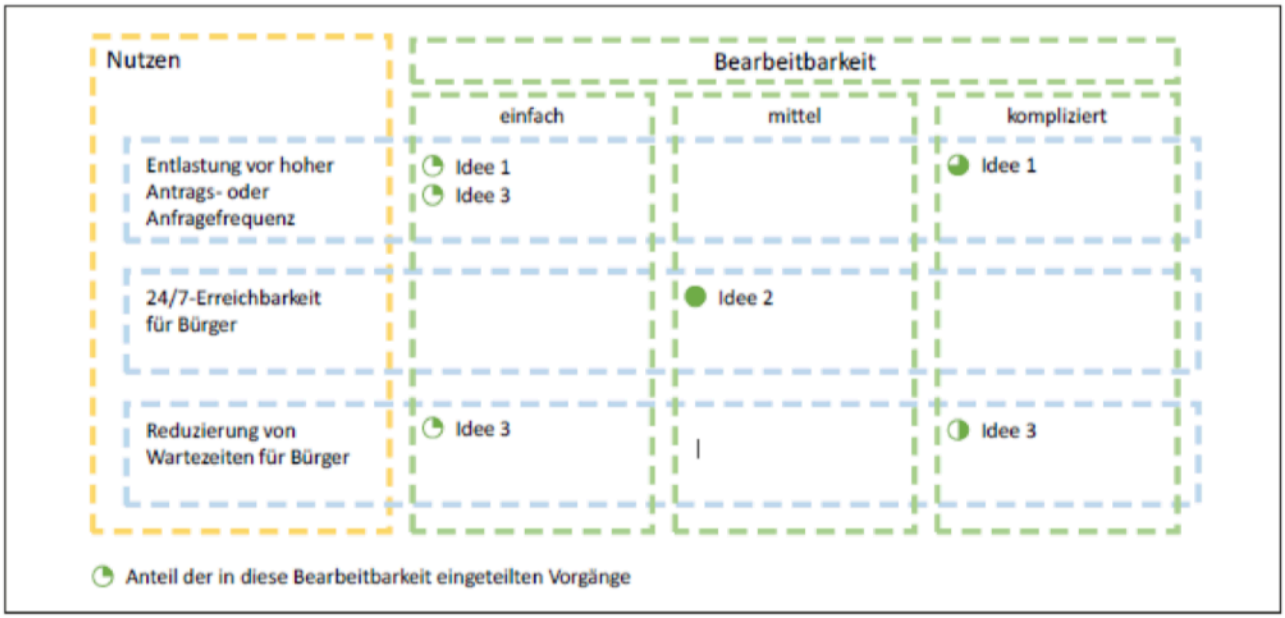


Abb. 1: Beispiel für ein einfaches Auswahlschema von Ideen zum Einsatz von KI [HV20, S. 38]

Der erfolgreiche Einsatz von KI benötigt eine Verortung in der Organisationsstruktur. Daher werden Vorschläge zur branchenspezifischen Einführung von KI in Kommunalverwaltungen unterbreitet, die sowohl strategische Überlegungen als auch die organisatorische Etablierung und Steuerung einer KI-Einheit wie notwendige Personalressourcen, Fachexpertise, Querschnittsdenken, organisationsübergreifende Koordination, und mögliche Formen interner und externer Kollaboration,

bspw. in Labs unter Einbezug der Stadtgesellschaft, als interkommunale Zusammenarbeit, in Kooperation mit Forschungsinstituten, ansprechen. Experimentelles, paralleles und iteratives Vorgehen wird dabei ebenso adressiert wie die Notwendigkeit von Partizipation und begleitendem Lernen. Abbildung 2 gibt Anregungen, zeigt Möglichkeiten und eröffnet Handlungsräume.

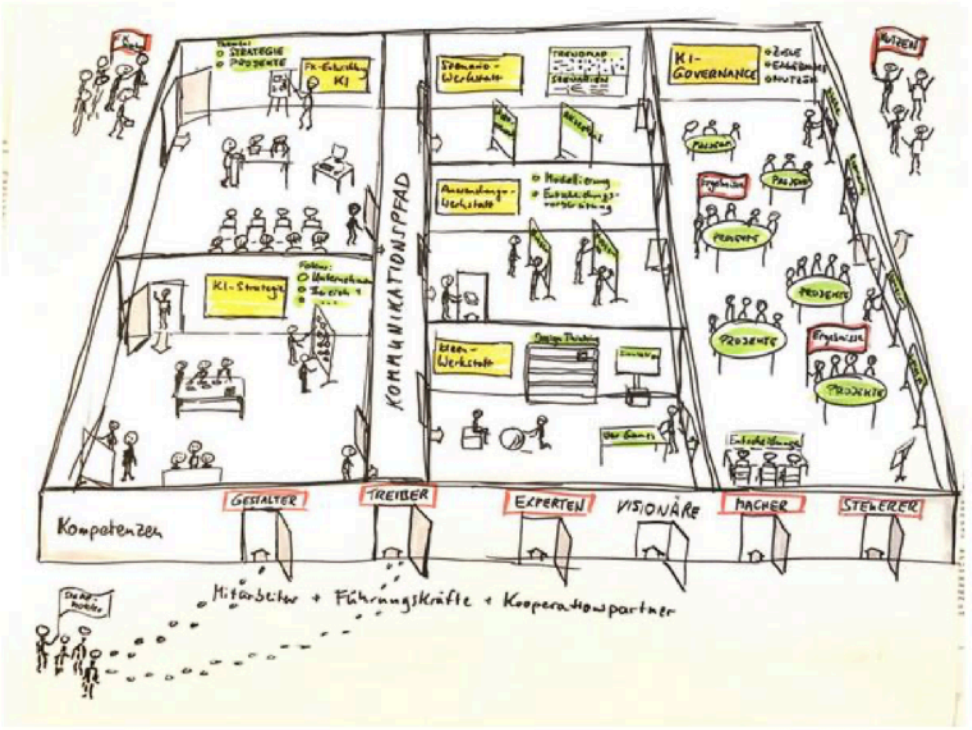


Abb. 2: Bausteine eines möglichen Vorgehens zur Einführung von KI [AI20]

Praxisbeiträge von Verantwortlichen aus Kommunen, die bereits erste Ansätze implementiert haben, um KI zum Wohl der Stadtgesellschaft zu nutzen, ergänzen die Publikation bewusst an zentralen Stellen. Eine ausführliche Literaturliste und ein Glossar beschließen die Ausführungen, so dass Gelegenheit zur Vertiefung des Themas gegeben und gleichzeitig ein gemeinsames Verständnis von KI möglich wird.

4 Ausblick

KI benötigt Aufmerksamkeit in den Kommunalverwaltungen und bietet vielfältige Chancen. KI ist aber mehr als nur „neue Technik", sie wird die Arbeitsteilung zwischen Mensch und Maschine qualitativ wie quantitativ verändern. KI-Projekte erfordern ein Umdenken im Umgang mit Daten, sie stellen andere Fragen an Vergabe, Projektmanagement, Betrieb, Organisation und Anspruchsgruppen als klassische IT und benötigen daher neue Antworten. Wissen, Fachkenntnis und persönliches Engagement der handelnden Akteure helfen nicht nur, diese Fragen zu beantworten, sondern können darüber hinaus vor Ängsten schützen, Widerständen entgegenwirken und zur Akzeptanz beitragen.

Die Publikation „Künstliche Intelligenz für die Smart City- Handlungsimpulse für die kommunale Praxis" ist die erste ihrer Art im deutschsprachigen Raum, die KI für den kommunalen Bereich kompakt erschließt und dessen spezielle Bedingungen zielgruppenspezifisch berücksichtigt. In Feedbacks zu Vorträgen, Podcasts und Online-Veröffentlichungen der vergangenen Monate ist deutlich geworden, dass die Publikation auf eine rege Nachfrage trifft, eine Wissenslücke füllt und bisher durchweg positive Resonanz und Rezensionen erhielt (s. u. a. [BS21] und [Me21]).

Ausblickend wäre eine empirische Erfassung der Informations- und Weiterbildungsbedarfe von Kommunalverwaltungen zum Thema KI, auch über die Vermittlung von Basiswissen hinaus, wünschenswert, die durch Umfragen und Experteninterviews realisiert und gestützt sein könnte. Die

Einfachheit der Darstellung der vorgestellten Publikation und die darin verortete Vermittlung von Basiswissen und Grundlagen zu KI könnten darüber hinaus auch für andere Branchen Vorbild sein und Nachahmung finden.

5 Literaturverzeichnis

[AI20] ai concepts: KI-Einstiegsmöglichkeiten. https://www.ai-concepts.com/ki-innovation.html Stand: 11.05.2021

[BS21] Behörden Spiegel: Lesenswert und praxisnah - “Künstliche Intelligenz für die Smart City“, ISSN 1437-8337, Nr. V / 37. Jg., 19. Woche, S. 22, 2021

[HV20] Hein, T.; Volkenandt, G.: Künstliche Intelligenz für die Smart City - Handlungsimpulse für die kommunale Praxis. Knowledge & Trends, Berlin 2020. https://www.ai-concepts.com/downloads/KI-fuer-die-Smart-City_E-Book-Ausgabe.pdf Stand: 11.05.2021

[Me21] Die intelligente Kommune. Kommune21, 22. April 2021. https://www.kommune21.de/meldung_36067_n.html Stand: 11.05.2021

Author information

Tabea Hein

Sie ist ausgebildete CDO, zertifizierte KI-Managerin und hat ihr Wirtschaftsinformatik-Studium mit einer Masterarbeit zu digitaler Transformation kommunaler Verwaltung mit Fokus auf den Einsatz von externem Consulting abgeschlossen. Sie ist Co-Autorin des Buchs „Künstliche Intelligenz für die Smart City“. Seit 2015 entwickelt sie als interne Beraterin in einer deutschen Großstadt Ideen zu Verwaltungsreformen und koordiniert die Maßnahmen.

Prof. Dr. Christian Schachtner

Er ist Studiengangleiter im Bereich Public Management der IU Internationale Hochschule und lehrt insb. in New Work, Öffentlichem Recht & Behördensteuerung, Public & Nonprofit Management sowie Nachhaltigkeitsmanage-ment. Daneben begleitet er als wissenschaftlicher Stabstellen-leiter für Prozessmanagement und Agilität und Inhouse-Consultant mit über 10-jährigen Führungserfahrung Digitalisierungsprojekte der öffentlichen Verwaltung. In verschiedenen Behörden fungiert er regelmäßig als selbstständiger Organisationsberater auf Bundes- und Landesebene bzw. im kommunalen Kontext.

Management der technischen Infrastruktur

AI Lab – An Ecosystem for Managing AI projects[1]

Martin Leucker[2], Maria Ostanina[3]

[1] This work was funded by the Bundesministerium für Bildung und Forschung (BMBF) through the KI-Lab Lübeck Project.

[2] Universität zu Luebeck, Institut für Softwaretechnik und Programmiersprachen, Ratzeburger Allee 160, 23562 Luebeck, Deutschland, leucker@isp.uni-luebeck.de

[3] Universität zu Luebeck, Institut für Softwaretechnik und Programmiersprachen, Ratzeburger Allee 160, 23562 Luebeck, Deutschland, maria.ostanina@isp.uni-luebeck.de

Abstract: This paper presents AI Lab, an ecosystem with tools and resources for building AI-based systems. It provides a basic infrastructure for building and implementing AI systems rapidly, with a focus on machine learning capabilities (challenges/ issues), i.e. training and running models, storing, analyzing, and processing data. More over, it supports the deployment of the built solutions on different target systems including a broad range of edge systems. The AI Lab is designed to serve both novices as well as experts in AI by knowledge transfer from university to practice, spanning from bachelors to researchers.

Keywords: artificial intelligence platform; integrated development platform

1 Introduction

The AI Lab Lübeck forms the basis of an ecosystem supporting the rapid development of machine learning-based solutions. While it supports educational purposes, it is especially suitable for research projects, and, in the long run, also for professional projects. It is built up by high-performance machine learning computers for high-end learning tasks as well as midsize servers for interactive development. Regarding software, it integrates the most popular open source software solutions for machine learning. Moreover, it offers a configurable deployment tool chain allowing to employ the learning results on the variety of edge devices. For smooth management of machine learning based projects it offers various collaboration capabilities, like cloud data storage and basic project management tools, an open data portal etc. The lab is accessible via ai-lab.digital-hub-luebeck.de. Figure 1 shows screen-shot of the main page of the lab.

The anticipated or realized functionality and design of AI Lab builds upon user studies performed with various project leaders of AI projects at the University of Luebeck and a study of similar labs like Google Colaboratory[1], Peltarion[2], Orange[3], Kaggle[4] respectively.

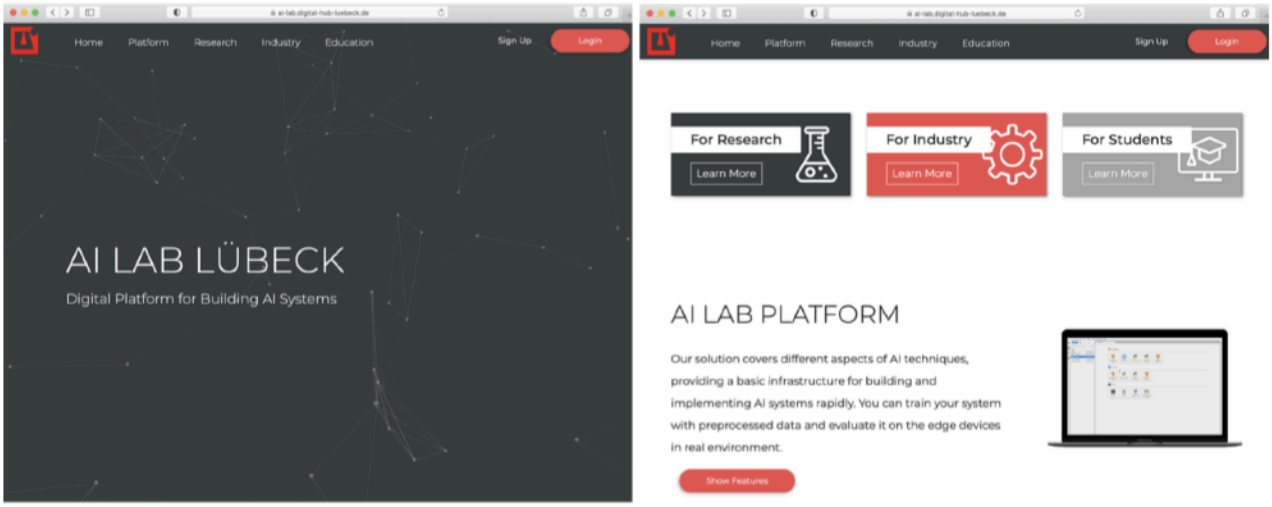


Fig. 1: Screenshots of the AI Lab

The AI Lab ecosystem grows within *KI-Lab Lübeck* project funded with almost 2 Million Euros by the German Ministry of Education and Research (BMBF).

2 The AI Lab from a Users Perspective

The AI Lab is designed to support users within machine learning projects. Such projects typically pursue the following workflow:

1. Data is collected and annotated, checked for quality and possibly enriched, to serve as the basis for machine learning.
2. Using specialized forms self probabilistic differential programming, suitable neural networks are trained and later tested.

[1] https://colab.research.google.com/notebooks/intro.ipynb

[2] https://peltarion.com/

[3] https://orangedatamining.com/

[4] https://colab.research.google.com/notebooks/intro.ipynb

3. If the trade model is considered to be adequate, it is deployed on one or more target platforms.

Currently, mostly image analysis as well as control solutions for robots are developed within the AI Lab.

Let us consider the anticipated workflow in more detail: The user either has or requires a login for the AI Lab platform. This login serves as single-sign-on and allows transparent access to each individual service according to the authorization provided by the AI Lab administrators.

The next step is typically to define a new project and to assign coworkers to the project, if needed. Each project leader can authorize project members to work on the project by defining access rights in a fine granular manner.

A project is composed of storage for data as well as algorithms, the latter, in our lab, typically python programs developed via Jupyter notebooks.

After uploading the corresponding data using the underlying Nextcloud facilities, it is analyzed using several data validation services available in the lab. The AI Lab comes with an open data platform which allows to use public data for the own project but also to share own data under several open data licenses.

Once the data is in reasonable shape,the user creates a Jupyter notebook and specifies the machine learning task. The user typically experiments with the program first on smaller machines but may schedule the program for batch processing on the high-performance computers, once the parameters for the learning program seem to be adequate. Nevertheless, the AI Lab supports persistent sessions allowing to suspend the work also within the interactive usage.

Figure 2 shows two screen-shots of using Juypiter notebook within the lab.

Fig. 2: Screenshots showing the use of the AI Lab

Once a model is successfully trained it may be tested regarding the quality and deployed to target systems. The lab provides the execution of Kubernetes Pods for using the trained model in the cloud but also transformation tools for running the solution on dedicated hardware. Via ONNX we support a variety of different edge platforms, ranging from Nvidia Jetson and similar Intel based solutions to dedicated FPGA boards (especially those from Xilinx).

Whenever several iterations of a typical machine learning workflow will be carried out, for example when continuously enriching the models based on new data, it is helpful to define individual learning workflows. Within the lab this is supported by kubeflow and elyra which provide a graphical user interface for defining complex machine learning workflows.

3 The AI Lab from a Technical Perspective

The AI Lab supports both the storage and the processing of data.

3.1 Storing of data

For storing of data, both a cloud space as well as an open data platform is provided. For the cloud space, AI Lab relies on Nextcloud, an open source cloud solution [Ne21]. Storage devices can be easily integrated into the Nextcloud via the webdav protocol, allowing to scale the data space continuously.

Moreover, we provide an open data portal which relies on the open source framework DKAN [DK21]. Based on the CMS Drupal, it provides both an API as well as a front-end to support batch processing and interactive usage, respectively.

Note that both, the Nextcloud as well as the open data portal (which also store data based on the webdav protocol), is limited in terms of latency and throughput. That means that for machine learning purposes, relevant subsets of the data to be considered should be copied to high-performance storage devices directly attached to machine learning devices (like the ones included in the NVIDIA DGX systems).

3.2 Processing of data

The AI Lab is build with scalability and resilience in mind. For processing of data, the AI Lab provides a variety of different hardware and software solutions. For the hardware part, high-performance systems from NVIDIA (one DGX-2 and one DGX-A100 to be extended by two further DGX-A100) for intensive machine learning jobs extend a set of smaller servers (DELL with GPUs) mainly used for solution engineering and small scale applications. All machines run jupyter notebooks [Ju21] to simplify interactive development. Moreover, we use kubeflow [Ku21] and elyra [El21]to support complex machine learning pipelines, persistent sessions and scheduling.

As software, a set of open source machine learning frameworks such as pytorch [Py21] and tensorflow [Te21] are provided in the lab. The rapid development of new machine learning solutions will be reflected by a constant update and extension of the software frameworks available in AI Lab.

Moreover, the AI Lab supports the development for edge devices. It offers deployment tool chains implementing the ONNX standard [ON21] that enables AI developers to use models with a variety of frameworks, tools, runtimes and compilers.

The overall platform is enriched with monitoring and statistics capabilities to ensure the quality of service of the overall system.

4 Conclusion and Future Work

In this paper, we briefly described the AI Lab currently built up in Lübeck. It supports the rapid and continuous development of machine learning based systems. In the future, we plan to enrich both the hardware of our system but especially also the engineering process of such systems by further methods and methodologies. The AI legislation [EU21] currently under discussion in the EU will require new engineering process support especially for high risk AI-based systems and it is our goal to provide such support in the AI lab in the new future.

References

[DK21] DKAN Community: DKAN Documentation, https://demo.getdkan.org/modules/contrib/dkan/docs/index.html modules/contrib/dkan/docs/index.html, Accessed 2021-05-15, 2021.

[El21] Elyra Team: Elyra Documentation, https://elyra.readthedocs.io/en/latest/getting_started/overview.html, Accessed 2021-05-15, 2021.

[EU21] EU Commission: Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence, https: //ec.europa.eu/commission/presscorner/detail/en/IP_21_1682, Accessed 2021-05-15, 2021.

[Ju21] Jupiter Team: Jupiter Documentation, https://jupyter.readthedocs.io/en/ latest/, Accessed 2021-05-15, 2021.

[Ku21] Kubeflow Authors: Kubeflow Documentation, https://www.kubeflow.org/ docs/, Accessed 2021-05-15, 2021.

[Ne21] Nextcloud: Nextcloud Solution Architecture: Bring data back under control of IT, http://nextcloud.com, Accessed 2021-05-15, 2021.

[ON21] ONNX Team: ONNX Documentation, https://onnx.ai/, Accessed 2021-05-15, 2021.

[Py21] PyTorch Team: PyTorch Documentation, https://pytorch.org/docs/stable/index.html, Accessed 2021-05-15, 2021.

[Te21] TensorFlow Team: TensorFlow Documentation, https://www.tensorflow.org/guide, Accessed 2021-05-15, 2021.

Author information

Prof. Dr. Martin Leucker

Martin Leucker is currently a professor at the University of Lübeck, Germany heading the Institute of Software Engineering and Programming Languages. He obtained his Ph. D. at the RWTH Aachen, Germany and afterwards, he worked as a Postdoc at the University of Philadelphia, USA and at the Uppsala University, Sweden. He pursued his habilitation at the TU München, Germany. He is the author of more than 100 peer reviewed conference and journal papers ranging over software engineering, formal methods and theoretical computer science.

Maria Ostanina

Maria Ostanina is a member of the scientific staff at the Institute for Software Engineering and Programming Languages at the University of Lübeck. She is the project manager of the "KI-LAB Lübeck" project, focusing on building an AI infrastructure and co-developing new educational programs and qualification activities. Her research activities are concerned with building and managing integrative collaborative platforms.

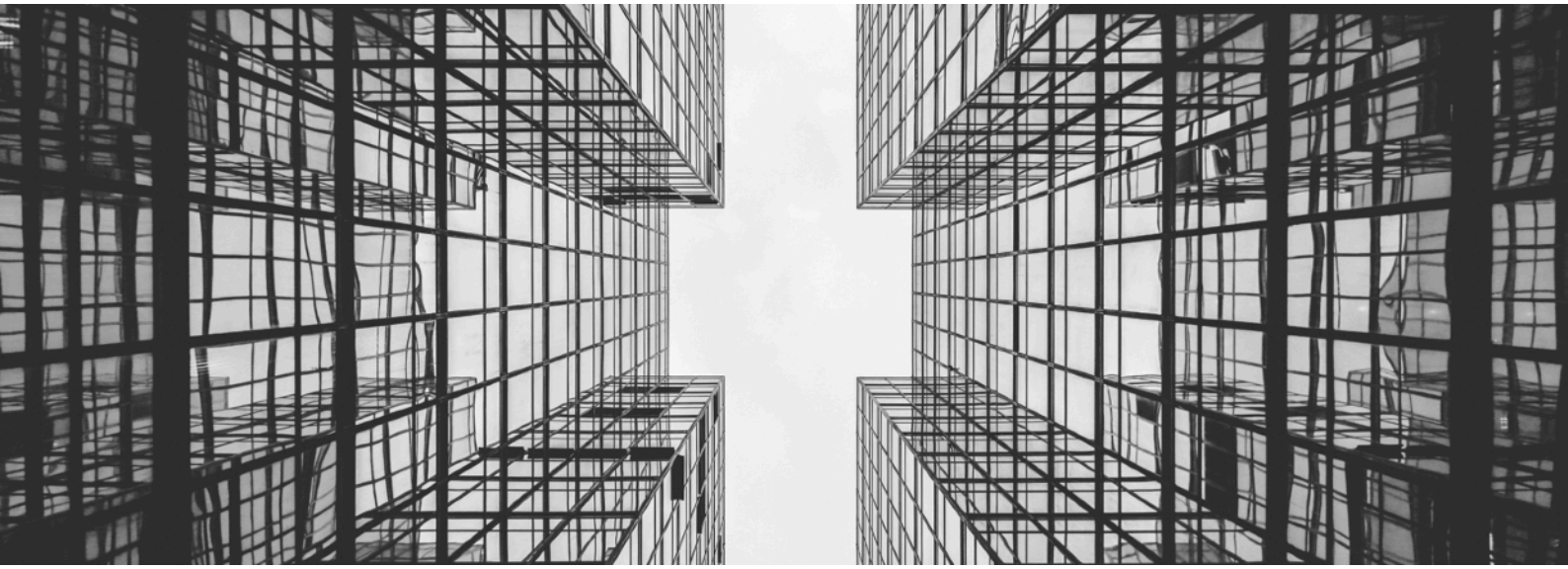
Towards Reference Architectures for Artificial Intelligence: A Preliminary Report

Alexis T. Bernhard[1], Jana Koehler[2], Sophia Saller[3]

[1] Saarland University, Saarland Informatics Campus E1.1, D-66123 Saarbruecken, Germany alexis.bernhard@uni-saarland.de

[2] Deutsches Forschungszentrum fuer Kuenstliche Intelligenz and Saarland University, Saarland Informatics Campus D3 2, D-66123 Saarbruecken, Germany jana.koehler@dfki.de

[3] Deutsches Forschungszentrum fuer Kuenstliche Intelligenz, Saarland Informatics Campus D3 2, D-66123 Saarbruecken, Germany sophia.saller@dfki.de



Abstract: Over the last decade, companies increasingly launched initiatives to adopt Artificial Intelligence (AI) technologies, notably machine learning, for innovation and automation. Many of these initiatives involve complex software projects, which implement the technology. Whereas the AI core technologies come as ready-to-use packages, which matured over the last years, the solution architecture of AI application systems is developed from scratch in these projects due to the lack of reference architectures for common AI application scenarios. In this paper, we review existing architectural styles, which have been proposed or are relevant for AI systems such as the lambda architecture, the blackboard architecture, and agent architectures, and discuss their degree of maturity and possible refinements.

Keywords: Artificial Intelligence (AI) architectural styles; reference architecture; Lambda architecture; Blackboard architecture; Agent Architectures

1 Architectural Styles for Artificial Intelligence Systems

Over the last decade, companies increasingly adopted Artificial Intelligence (AI) technologies, notably machine learning, in their application systems, see for example [Ra20, Fig. 2-2]. The learning curve to build these systems is steep and many software projects fail to reach their goals as effort and costs are often underestimated. Whereas AI core technologies come as ready-to-use packages such as Tensorflow or cloud services on AWS, Google, or other platforms, the solution architecture of AI application systems needs to be developed from scratch in these projects due to the lack of reference architectures for common AI application scenarios.

In this paper, we review relevant architectural styles that have been proposed in the context of AI application systems and discuss their maturity as an AI reference architecture. Reference architectures combine general architecture knowledge and general experience with specific requirements for a coherent architectural solution for a specific problem domain [NOB11; Vo11]. As experience in AI solution creation is growing, it is possible to generalize the many from-scratch built architectural solutions into reusable reference architectures, which help in structuring AI application systems based on proven patterns and styles and thus reducing the effort, costs, and risks associated with AI application projects.

A relevant architectural style, which emerged in the context of Big Data applications, is the **Lambda architecture**, see Fig. 1. Lambda is a scalable and fault-tolerant data-processing architecture for batch- and streaming data [Ki15; MW13; Se16]. As many AI applications use data-intensive machine learning (ML) technologies, Lambda is important for AI applications to organize the data pipeline for training and executing ML components. The Lambda architecture consists of three layers, a *data consumption* layer, which integrates various data sources, a *data processing* layer, which is split into two parallel layers for stream and batch data, and a *presentation & serving* layer, which provides data results to end users and consuming applications. The layers support 6 phases of working with the data where 3 of them focus on preparing data for a later analysis. The analysis phase in Lambda envisions to integrate AI solutions for data analysis, however, does not provide any further details. Here, a great need for further refinement is obvious.

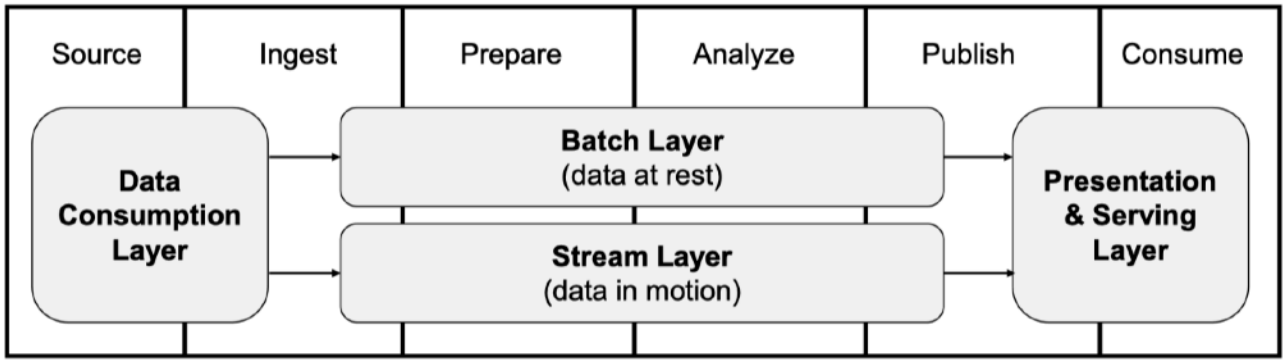


Fig. 1: The Lambda architecture based on [MW13; Se16].

The **Blackboard architecture** has been proposed for the integration of intelligent agents (AI systems) that cooperatively work on solving a problem [CL94; Er80; Sc13], see Fig. 2. A blackboard architecture comprises three key components: the *Blackboard* itself as an accessible data store and integration solution, from which intelligent agents read data and to which they write their computational results, the *agents*, which are autonomous AI applications, and the *controller*, which controls when and how different agents can access the blackboard. At this level of description, the blackboard architecture is a very generic solution for cooperative problem solving by AI systems. Unfortunately, concrete implementations show only very application-specific solutions, which are hardly reusable in other contexts. There is a strong need to refine the blackboard architectural style into more concrete, but still reusable components, which includes to more precisely define the interfaces and control- and data-flow between the blackboard, controller, and agents.

Finally, AI research has introduced several **agent architectures**. Modern AI research is based on the metaphor of the intelligent agent [PW05; RN09], which is capable of perceiving information from an environment through its sensors, uses this information to take decisions, and then selects actions to act in the environment based on the taken decision. A selected action is executed using the actuators of the agent. The most famous and widely used AI textbook [RN09] introduces the metaphor of the intelligent agent as well as architectures for 5 different types of agents such as the simple reflex agent, the model-based reflex agent, the goal-based, utility-based, and the learning agent. For modern AI systems, the architecture of the learning agent is relevant for any ML system in its training phase. The utility-based agent, see Fig. 3 is relevant for any AI/ML at work. The learning agent is rarely seen in practice today, as it assumes an AI system that pursues its own goals. In the remainder of this paper, we focus on the utility-based agent architecture.

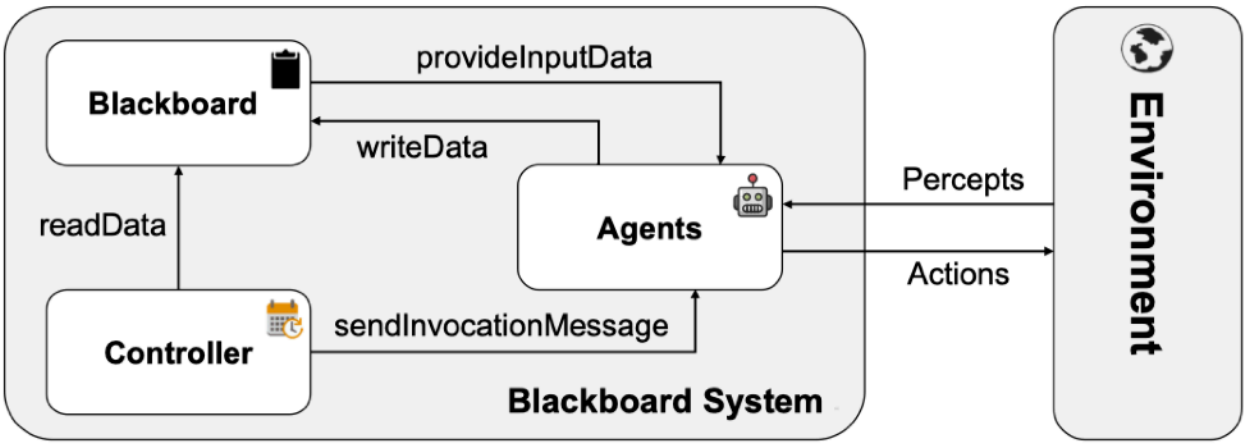


Fig. 2: The blackboard style based on [Ni86].

The architecture shown in Fig. 3 displays eight other model elements, mostly formulated as questions, besides the *sensors*, through which the agent perceives the environment, and the *actuators*, through which the agent performs actions. The first question, which needs to be answered based on the data collected by the sensors, is *What the world is like now?* and refers to the capability of an AI agent to assess the state of the environment. To do so, the agent needs to update its internal *state* as the internal representation of the current situation in the environment. To perform the state update, the agent needs to answer the questions of *How the world evolves?* and *What my actions do?*, which means that it needs to have access to an internal causal model, which allows it to predict changes in the world as well as the consequences of its own actions. The combined answers result in an updated internal world model, based on which the agent can answer the question *What it will be like, if I do action A*. To evaluate several possible and alternative actions, an agent needs a utility model to assess the desirability of a change in the world caused by its action and to answer the question of *How happy I will be in such a state?* The agent scores the state resulting from a world change using the utility model. The action that maximizes the utility of the new state, based on the anticipated action outcome, is selected by the agent for execution and sent to the actuators.

The architectures as shown in the book by Russell and Norvig hardly meet the quality standards of proper architectural models for reference architectures. They introduce some components, however, the interfaces between these components are not properly defined and the semantics of the arrows is rather unclear. The dashed arrow in the upper part of the figure depicts the update of the world model, the left hand-side boxes show internal models used by the agent, and the right-hand side boxes show the flow of reasoning within the agent. A reasoning component is depicted by a question and the arrow from a left-hand side box to a right-hand side box shows that in order to answer the question, the agent needs to apply the knowledge stored in one of the models. As in the case of the Blackboard architecture, the agent architecture for the utility-based agent as well as all other agent architectures need further clarification and refinement, which we discuss in the next

section. Furthermore, technology mappings and reusable reference implementations must be worked out on the route towards proper reference architectures.

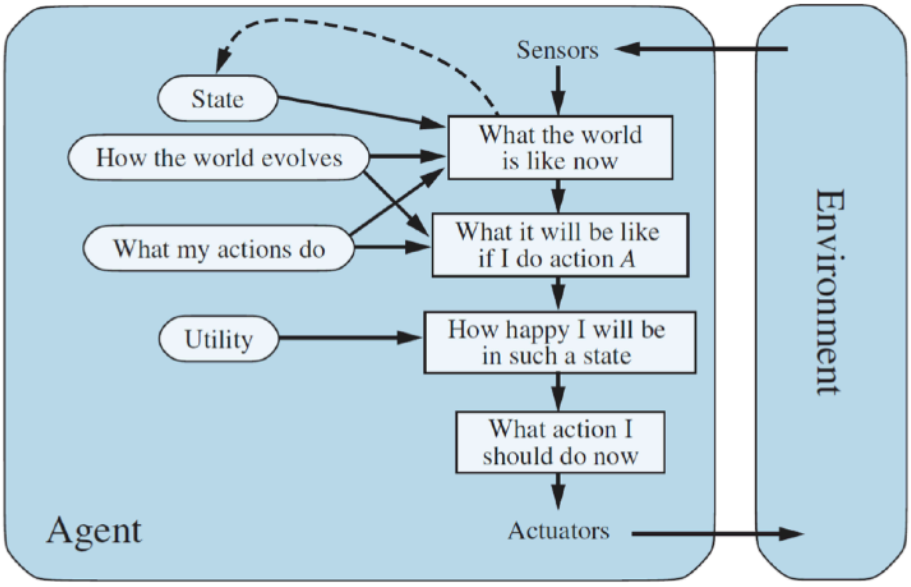


Fig. 3: The utility-based agent by [RN09].

2 Clarification and Refinements of the Utility-based Agent Architecture

The agent architecture by Russell and Norvig gives an overview over the models and reasoning process behind a utility-based agent. In the following, we propose a clarification of this architecture towards an architectural style as shown in Fig. 4 and then show how this architectural style can be further refined for a specific class of AI application systems.

Following the data flow in the utility-based agent, the first change is a *Data Manager* component, which combines the state model element with the *What the world is like now?* reasoning component. The data manager component combines the data of all sensors and the historical data in the state model to compute an updated model of the current state of the environment. The reasoning components to answer the questions of *How the world evolves?* and *What my actions do?* are combined into the *Model* component where the causal background knowledge of the agent is stored. The question *What it will be like, if I do action A?* is answered by a new *Simulation* component to simulate the outcome of a possible candidate action. The *Evaluation* component uses the utility to answer the question *How happy I will be in such a state?* and decides which action to take.

The architectural style shown in Fig. 4 makes components and their purpose more clear and is more concrete with respect to the data flow between the components. In the following, we discuss how this architectural style is refined to provide an architecture for two specific application systems, which are typical representatives of two practically relevant classes of AI applications: natural-language-based information retrieval and intelligent optimization and control.

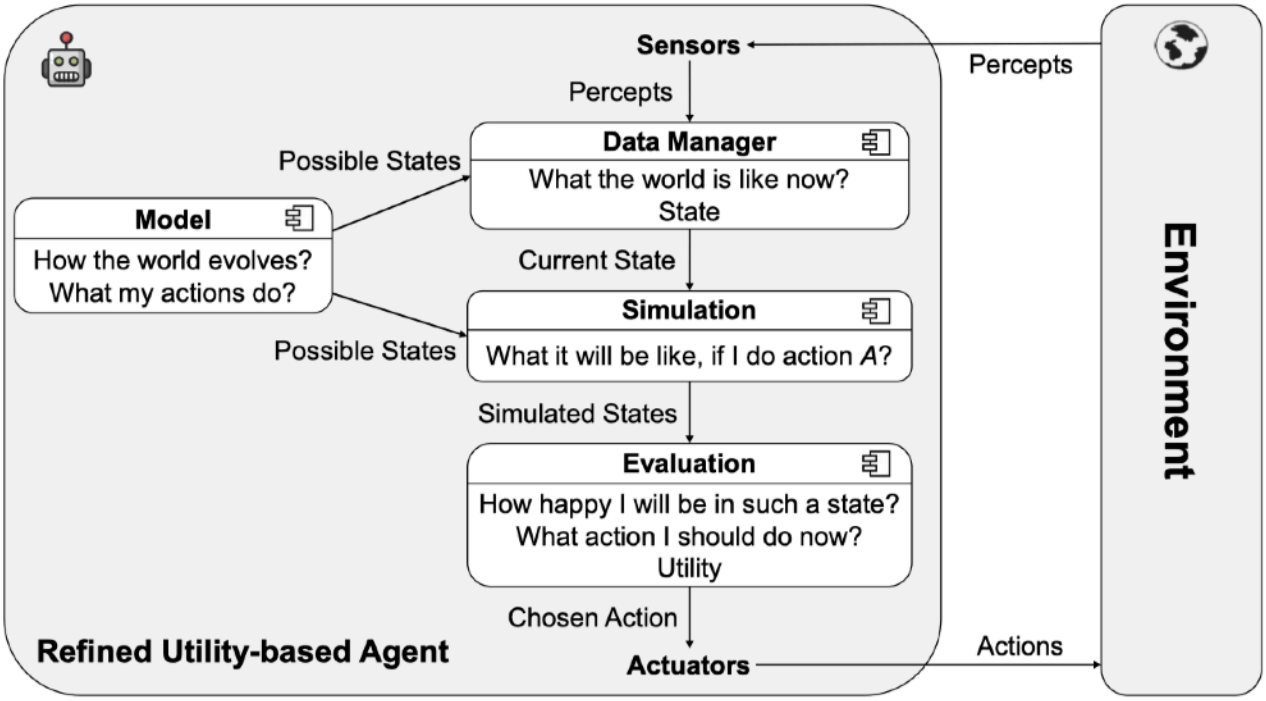


Fig. 4: The clarified utility-based agent architectural style.

2.1 AIden: A Natural-Language-based Information Retrieval Agent

`AIden` is an information retrieval [SMR08] agent used in e-learning to answer questions from students about content taught in the AI lecture at Saarland University. Students enter questions as text over a web-based interface and listen to the explanation of retrieved relevant slides given in the lecture. Fig. 5 shows the components of `AIden` as a refinement of related components in the clarified utility-based agent architecture.

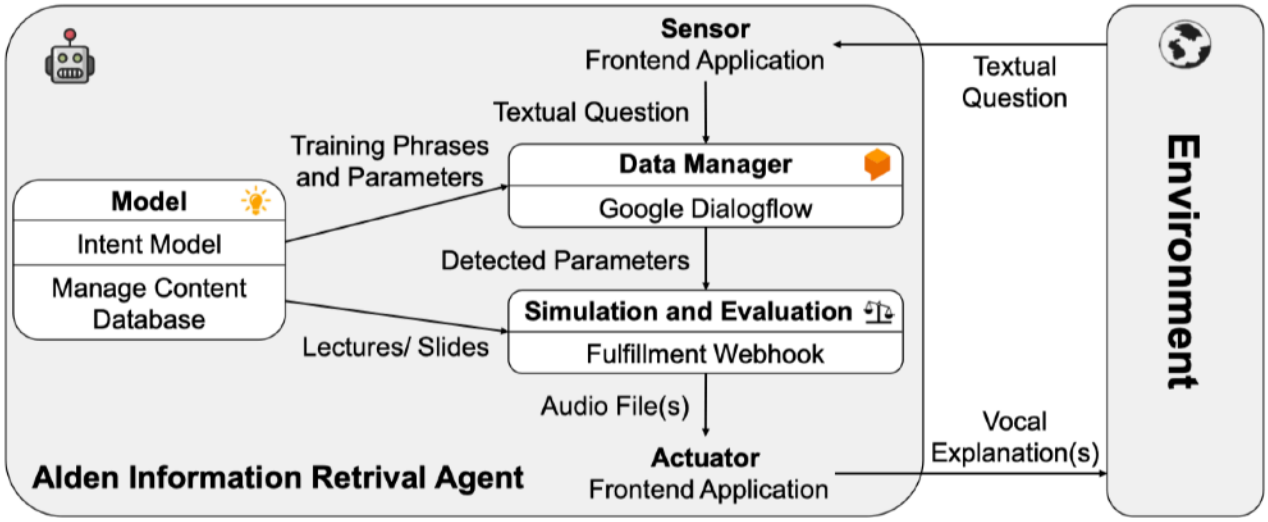


Fig. 5: The clarified utility-based agent architectural style.

The sensor in `AIden` is a keyboard listener in the *Frontend Application* component, which also integrates the actuator to play the resulting audio file. The model consists of a *Manage Content* component containing annotated lecture slides and audio files generated using the Google TTS Engine [Go21a], and an *Intent Model*, which is trained to recognize the questions of students using Google Dialogflow [Go21b; SA20]. The Data Manager is based on Google Dialogflow, a natural

language understanding platform for the development of conversational interfaces. The component detects the intent in a given user request, which allows `AIden` to appropriately react to a user request. The simulation and evaluation components are combined in this system and implemented in a *Fulfillment Webhook* component, which computes the intent-based response of the system using common AI algorithms for document retrieval. Document similarity measures such as for example TF-IDF [SMR08] serve as the utility function.

2.2 CTWire: An AI Scheduling System using Constraint Optimization

CTWire [Ko20] is an AI scheduling system to produce the control for a Zeta machine by Komax AG, Switzerland, which manufactures cable trees. For these machines, an optimal wiring sequence of cable plugging operations has to be computed, which is to be followed by the machine to minimize production time and maximize production quality. The architecture of CTWire is shown in Fig. 6 as another refinement of the utility-based agent.

To obtain an optimal wiring sequence, the user inputs a description of the cable tree and its layout on the assembly palette into the *Frontend Application*, which serves as the sensor. The description and layout are passed as an XML file to the *Constraint Mapper*, which refines the Data Manager, and extracts constraint parameters from the XML input. The *Constraint Model*, together with the objective function, forms the model element. The objective function for the optimization refines the utility function. The *Constraint Programming Solver (CP Solver)* refines both the simulation and evaluation element and can for example be implemented using the IBM CPLEX CP solver [IB21]. Within the CP Solver, constraints are initialized by applying the constraint parameters from the Data Manager to the constraint descriptions from the Model. By simulating various wiring sequences and evaluating their utility based on the objective function, the CP Solver finds the optimal wiring sequence for the machine, which is passed on for execution to the Zeta machine acting as the Actuator.

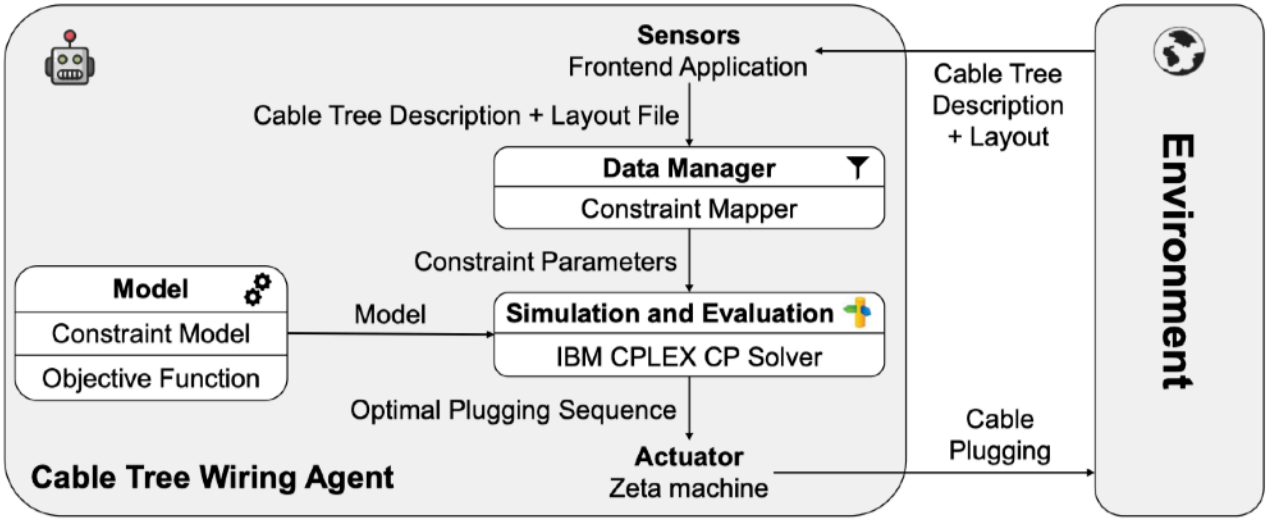


Fig. 6: Scheduling agent CTWire as a refinement of the utility-based agent.

3 Conclusion and Outlook

We reviewed the Lambda, Blackboard, and Agent architectures, which are relevant when implementing AI solutions. Whereas these three architectures introduce a generic system structure, many details are still missing to count them as proper AI architectural styles. As an example, we discussed a clarification of the utility-based agent architecture and two refinements representing important classes of AI application systems: natural-language based information retrieval and intelligent optimization and control. Future work includes to further detail these architectural styles, refine existing architectural patterns and tactics for AI applications, and add technology mappings as well as reference implementations.

References

[CL94] Carver, N.; Lesser, V.: Evolution of blackboard control architectures. Expert Systems with Applications 7/1, pp. 1–30, 1994, issn: 0957-4174, url: https://www.sciencedirect.com/science/article/pii/095741749490023X.

[Er80] Erman, L. D.; Hayes-Roth, F.; Lesser, V. R.; Reddy, D. R.: The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty. ACM Comput. Surv. 12/2, pp. 213–253, June 1980, issn: 0360-0300, url: https://doi.org/10.1145/356810.356816.

[Go21a] Google: Cloud Text-to-Speech basics, 2021, url: https://cloud.google.com/text-to-speech/docs/basics, visited on: 05/05/2021.

[Go21b] Google: Dialogflow ES basics, 2021, url: https://cloud.google.com/dialogflow/es/docs/basics, visited on: 05/05/2021.

[IB21] IBM: Cplex Optimization Studio, 2021, url: https://www.ibm.com/products/ilog-cplex-optimization-studio/, visited on: 05/21/2021.

[Ki15] Kiran, M.; Murphy, P.; Monga, I.; Dugan, J.; Baveja, S. S.: Lambda architecture for cost-effective batch and speed big data processing. In: 2015 IEEE International Conference on Big Data (Big Data). Pp. 2785–2792, 2015.

[Ko20] Koehler, J.; Bürgler, J.; Fontana, U.; Fux, E.; Herzog, F.; Pouly, M.; Saller, S.; Salyaeva, A.; Scheiblechner, P.; Waelti, K.: Cable Tree Wiring–Benchmarking Solvers on a Real-World Scheduling Problem with a Variety of Precedence Constraints. arXiv preprint arXiv:2011.12862/, 2020.

[MW13] Marz, N.; Warren, J.: Big Data: Principles and best practices of scalable real-time data systems. Manning, 2013.

[Ni86] Nii, H. P.: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures. AI Magazine 7/2, p. 38, June 1986, url: https://ojs.aaai.org/index.php/aimagazine/article/view/537.

[NOB11] Nakagawa, E. Y.; Oliveira Antonino, P.; Becker, M.: Reference Architecture and Product Line Architecture: A Subtle But Critical Difference. In (Crnkovic, I.; Gruhn, V.; Book, M., eds.): Software Architecture. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 207–211, 2011, isbn: 978-3-642-23798-0.

[PW05] Padgham, L.; Winikoff, M.: Developing intelligent agent systems: A practical guide. John Wiley & Sons, 2005.

[Ra20] Rammer, C.; Bertschek, I.; Schuck, B.; Demary, V.; Goecke, H.: Einsatz von Künstlicher Intelligenz in der Deutschen Wirtschaft: Stand der KI-Nutzung im Jahr 2019, ger, ZEW-Gutachten und Forschungsberichte, Berlin, 2020, url: http://hdl.handle.net/10419/222374.

[RN09] Russell, S. J.; Norvig, P.: Artificial Intelligence: a modern approach. Pearson, 2009.

[SA20] Sabharwal, N.; Agrawal, A.: Introduction to Google Dialogflow. In: Cognitive Virtual Assistants Using Google Dialogflow: Develop Complex Cognitive Bots Using the Google Dialogflow Platform. Apress, Berkeley, CA, pp. 13–54, 2020, isbn: 978-1-4842-5741-8, url: https://doi.org/10.1007/978-1-48425741-8_2.

[Sc13] Schmidt, D. C.; Stal, M.; Rohnert, H.; Buschmann, F.: Pattern-oriented software architecture, patterns for concurrent and networked objects. John Wiley & Sons, 2013.

[Se16] Serra, J.: What is the Lambda Architecture?, 2016, url: https://www.jamesserra.com/archive/2016/08/what-is-the-lambda-architecture/, visited on: 05/27/2021.

[SMR08] Schütze, H.; Manning, C. D.; Raghavan, P.: Introduction to information retrieval. Cambridge University Press Cambridge, 2008.

[Vo11] Vogel, O.; Arnold, I.; Chughtai, A.; Kehrer, T.: Software architecture: a comprehensive framework and guide for practitioners. Springer Science & Business Media, 2011.

Author information

M. Sc. Alexis T. Bernhard

Er ist seit 2020 wissenschaftlicher Mitarbeiter und Doktorand am Lehrstuhl an der Universität des Saarlandes. Im Forschungsbereich Analyse und Optimierung von Prozessen arbeitet er an einer Dissertation im Themengebiet der erweiterbaren Architekturen und Projektmanagement für künstliche Intelligenz. Er schloss 2020 sein Studium mit einem Master of Science in Informatik am Karlsruher Institut für Technology (KIT) ab. 2016-2017 absolvierte er ein akademisches Jahr am University College Cork (UCC) in Irland und 2019-2020 schrieb er seine Masterarbeit in Kooperation mit der Linnéuniversitet in Växjö.

Prof. Dr. Jana Koehler

Sie ist seit 2019 wissenschaftliche Direktorin des Forschungsbereiches Analyse und Optimierung von Prozessen am Deutschen Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) und Inhaberin des Lehrstuhls für Künstliche Intelligenz an der Universität des Saarlandes. Sie hat Informatik und Wissenschaftstheorie an der Humboldt Universität zu Berlin studiert und 1994 an der Universität des Saarlandes in Informatik promoviert. Weitere berufliche Stationen waren die Albert-Ludwigs-Universität Freiburg, wo sie 1996 habilitierte, die Schindler AG, IBM Research Zürich und die Hochschule Luzern. Ihr Spezialgebiet sind KI-Methoden für die flexible Optimierung und Entscheidungsunterstützung in Fertigungs- und Geschäftsprozessen.

Dr. Sophia Saller

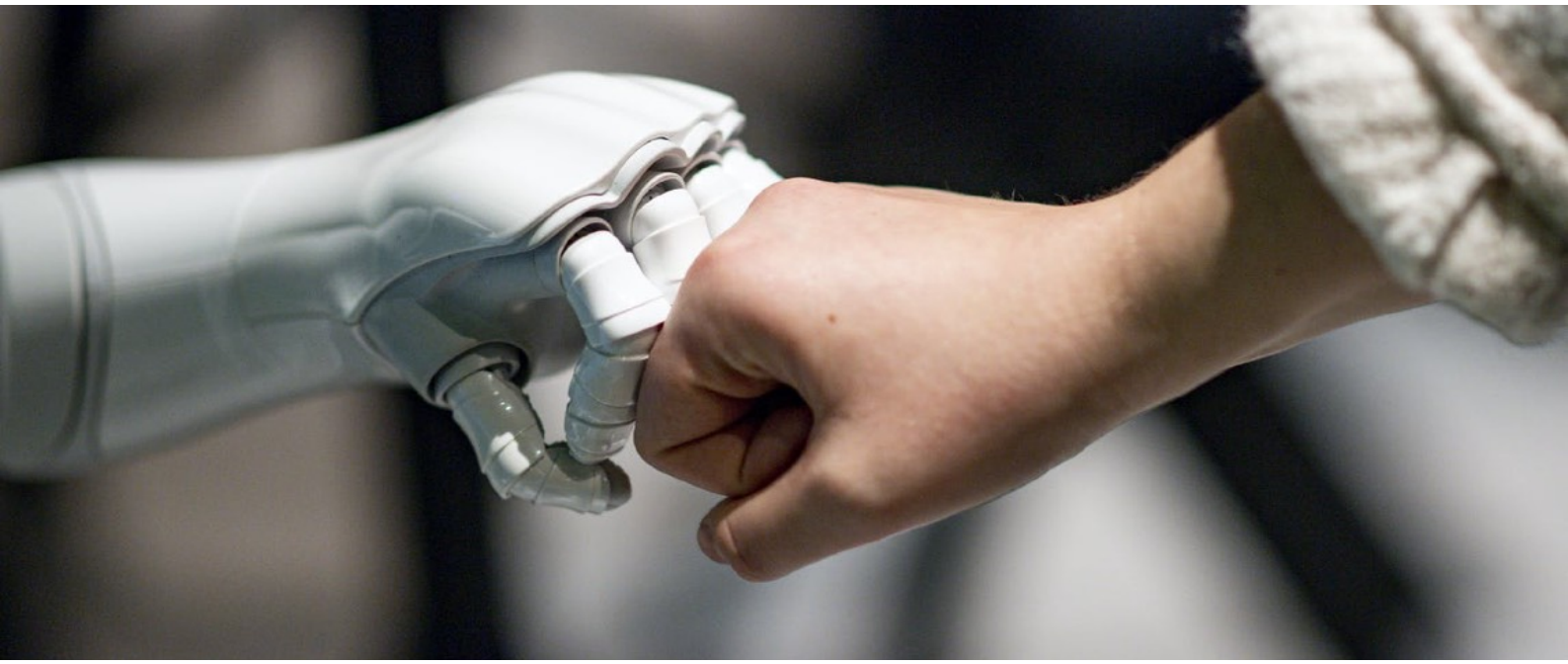
Sie ist seit 2020 wissenschaftliche Mitarbeiterin und Postdoc am Deutschen Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) und Forschungsbereichsmanagerin des Forschungsbereiches Analyse und Optimierung von Prozessen am DFKI. Von 2011 bis 2015 studierte sie Mathematik mit „first class honours“ an der Universität Oxford. Sie promovierte 2019 in Oxford mit einem Thema im Bereich Kombinatorik und Graphentheorie. Während dieser Zeit war sie zudem Mitglied der Deutschen Triathlon Nationalmannschaft und erhielt zahlreiche renommierte Sportstipendien. Im Jahr 2014 wurde sie Triathlon-U23 Weltmeisterin und gewann von 2012 bis zum Ende ihrer aktiven Zeit 2019 weitere zahlreiche Auszeichnungen. In ihrer aktuellen Forschungsarbeit beschäftigt sie sich mit neuen Methoden für die Modellierung im Bereich Constraint Programming.

AI-ware: Bridging AI and Software Engineering for responsible and sustainable intelligent artefacts

Dagmar Monett[1], Claudia Lemke[2]

[1] Berlin School of Economics and Law, Computer Science Department, Alt-Friedrichsfelde 60, Berlin, 10315, dagmar.monett-diaz@hwr-berlin.de

[2] Berlin School of Economics and Law, Information Systems Department, Alt-Friedrichsfelde 60, Berlin, 10315, claudia.lemke@hwr-berlin.de



Abstract: The Computer Science fields Artificial Intelligence (AI) and Software Engineering need to strengthen at their interception. This paper contributes to it by defining what is meant by AI concretely (spoiler: it is not only machine learning), by presenting related work that builds upon bridging the above mentioned fields, and by suggesting the AI-ware approach for paving the path to the intelligent artefacts of the future.

Keywords: Artificial Intelligence, Software Engineering, AI-ware, AI ethics, machinelearnization fallacy

1 The AI background

1.1 What is AI?

Historically, there has been a typical and profound misunderstanding of what machine or artificial intelligence (AI) is or what it can do. This has been chasing the advancements in the AI field thereby lacerating its credibility since its foundation [Mc76; Mi19; Wo20]. Promises of intelligent capabilities in machines have been pervasive in the AI field for decades, especially during the "AI Springs" or periods of time where (perceived) significant results are achieved. Enthusiasm grows, consequently, as do expectations. New developments and the field itself are "ballooned to an importance beyond [their] actual accomplishments" [Sh56] and it does not take long until disappointment arises, funding is cut, and interest in the field decreases.[1] AI remains far from a magic bullet that can solve all our problems.

Many are the reasons for those AI highs and lows. Perhaps one of the most significant ones is not having consensus on defining AI, both as a field and as its most important concept [MLT20; Wa19]. Understandable, there will always be subfields that flourish more or less depending on their success; other new ones originate and others merge. But AI silos and corresponding ivory towers are characteristic of the inability to reach consensus on AI's most fundamental concepts and techniques. Few has changed since 1956, when the field is said to be founded.

Most recently, the "machinelearnization fallacy," as we define and refer to it in our approach, has dominated public and practitioners' perception of AI. We argue that this fallacy is the urge to reduce reality to models and treat all problems by applying machine learning (ML), i.e. data-driven statistical-based algorithms (like artificial –deep– neural networks and related techniques), even when they are not the most suitable ones. The machinelearnization fallacy appears when it is thought that ML algorithms are the only ones that can solve any problem in any domain or lead to truly intelligent systems because they have already had success optimizing certain problems in very narrow domains. The implications of misclassifying cats or objects in an image, for instance, should never be equalled to those impacting society and its structures. Moreover, the implications of referring to AI when it is actually meant only one of its subfields can be harmful [Cr21], also to the AI community and its progress [La21].

The incremental advances that have been made in AI are mainly in low-level pattern recognition skills, however [Jo19]. They are nowhere near advanced enough to be considered intelligent [MD19; Mi19; Sm19; Wo20]. In fact, the lately over-celebrated advances in the AI field are recording an unparalleled eclipsing of other AI subfields, simultaneously obscuring the harms and risks that their outcomes can pose. The hype is nothing but a product of loud promoted and corporate-backed achievements in a few AI areas. Consequently, the public and other stakeholders have been provided a narrow, partial view of AI. Yet, ML is one among many AI subfields, the one that is currently getting the most attention and investments, though.

1.2 Why now?

AI, the new star in the emerging technology sky, owes its popularity to various technological developments from recent decades. On the one hand, AI's technological progress has been possible due to an exponential growth in the performance of hardware components, the ongoing increase of the internet of things, as well as the high computing and storage capacities through cloud computing and improved algorithms. Furthermore, open source technology has facilitated the rapid spread of algorithms' code and experiments, allowing for new implementations and applications. On

[1] These other periods are known as "AI Winters."

the other hand, the Big Tech's drive for collecting and analysing consumer data fosters a fairy-tale imagination of all established companies across industries to be able to emulate them. Techniques or approaches like big data offer new ways for business analytics and especially for using primary or very specific ML approaches to predict a certain future of these companies. It seems so easy on the surface, that the training and test purposes of AI algorithms using the apparently available own business data are aimed at guaranteeing complete and timely information business decisions. At the same time, AI- driven algorithms promote a further automation of operational processes and structures, which, in the end, secure the efficiency and effectiveness of the company in the medium to long term. Therefore, it is not surprising that AI projects are on the strategic agendas of company leaders, actually for years now, with high annual investments. Other factors like open access publications and peer reviewed literature on AI, available on preprint servers even before they are finally published, also offer an easy entry to the academic or scientific AI community, combined with open, online education about AI.

Together, these factors have transformed the power relations between consumers, producers, and prosumers of digital technologies, either based on AI or not. For example, it is now straightforward to learn about and produce AI applications with a minimum of effort and resources. Consumers of these technologies become also data sources, testers, or even producers of new applications themselves. Nevertheless, and despite the factors that fuel the current AI momentum, it is often alleged that the design and implementation of AI algorithms or AI-based systems must be subject to other "laws" distinct from the ones guiding the development of traditional IT in the field of Software Engineering. Which are these differences exactly? Are there any similarities? Could these two core areas of the Computer Science field learn from each other?

1.3 AI *is* software!

Actually, it all comes down to software, yet it is treated as if it were magic in the case of AI. A blurry line, which got thicker with time, was artificially produced between AI and Software Engineering, in particular, and Computer Science, in general. It does not come as a surprise that whilst Software Engineering evolved to the very established discipline that it is today, the AI field was stuck in algorithmic development, mainly improvised in academic settings with few practical applications at large scale. The traditional software structure that encompasses not only all software development phases but also its management and quality assurance [So16] is indeed different for some AI-enabled systems, but not all of them. They require more algorithmic tuning, uncertainty, ambiguity, a different handling of data (for those where it is a critical resource), and a higher margin of error and instability along the systems [HMO19; Oz20].

2 The AI-ware agenda

2.1 What does AI-ware mean and what is it not?

We define AI-ware as a term with a twofold meaning. On the one hand, we mean *the need to be aware of* what AI is and is not, its subfields, as well as a critical appraisal of where does applying AI make sense at all. Section 1 clarifies these aspects. On the other hand, we mean *the conjunction of* AI and Software Engineering, mainly –but not only– the former learning from the variety of well-established techniques and good practices from the latter, at the same time extending them. Additionally, AI-ware combines research movements and/or subfields from the Software Engineering and AI fields, like AI4SE and SE4AI [Mc20], but focused on an equal consideration more than a causal argumentation. AI-ware is neither a new subfield of Software Engineering nor a new form of experimenting with AI. Rather, it should take into account the specific features in the application of

the various AI methods and models, enable a targeted and economically arguable development of AI-based applications, and be understood as an emerging, modern field of Computer Science.

Software Engineering principles *do apply* to engineering AI-based systems [HMO19]. Together with good practices and better methodologies that are coming from some AI subfields [MBL20], they are paving the way to bridge both fields. Additionally, a shift from "model-centric AI" to "data-centric AI" is happening in some industries, along with the extension of DevOps to MLOps for the specific case of the ML pipeline. This resembles only part of the AI subfield, however, since not all problems need to be solved by applying ML (refer to the machinelearnization fallacy discussed above).

An AI-ware approach should bring the pieces of a complex puzzle together and open directions for further research, thereby

- acknowledging what AI is and which subfields, methods, and algorithms are part of it or come into question when a particular problem is to be solved;
- including the ethical stance in all of its phases, particularly when developing data- centric AI;
- applying Software Engineering principles, methodologies, and good practices, e.g. those particularly helpful for AI-based software development;
- making decisions regarding platforms, technologies, and architectures suitable for AI, as well as for managing resources including data, people, and processes.

2.2 What a possible research agenda could look like

Future AI applications for business environments need the design and development of responsible and sustainable intelligent artefacts [Di19]. There is often a gap between the idea that programmers have when solving a problem and the creative use of its solutions by users. Humans make techno-moral choices, which can have positive but also negative reinforcing effects. It is therefore essential that an AI-ware approach, primarily and above all, can close these gaps. One of the main niches for research action would be to consider how to integrate AI ethical challenges into classical Software Engineering frameworks. Thus, suggested research perspectives for an AI-ware agenda are:

- **Societal and ethical perspective** regarding the purpose of AI-driven applications, their impact on society and nature, and their contribution to a sustainable future. Further, regarding which problems urge to be solved and which AI applications need regulation because of their high risk to erode human rights, for instance.
- **Algorithmic perspective** regarding a realistic consideration of what kind of AI is needed, if at all, and for solving which problems. The machinelearnization fallacy, among other fallacies intrinsically related to the nature and function of AI algorithms, should be kept under criticism as long as other approaches (e.g. from symbolic AI, hybrid techniques, etc.) have not been evaluated. The solution to a problem does not necessarily require AI, let alone ML spectacles.
- **Data-oriented perspective** regarding the value of data, like property rights of data, its variety, and quality. It is no longer just the interesting user data but the big object-based data that allows for completely new power structures in the sense of measuring our complete world and surroundings once having access to this data.
- **Framework-based perspective** regarding both hardware and software architectural issues, potentially different when designing AI or classical IT software (compare DevOps to MLOps). Areas that have so far been considered in an undifferentiated way should probably be

separated, too. For example, the application of statistical methods in business analytics is also possible by using ML algorithms, but they still belong to the domain of data mining or data science and, thus, require other competences than the classical Software Engineering ones.

- **Economics perspective** regarding the efforts and resources when designing AI- based applications, the entire software process, and the corresponding operational tasks. Examples show time and again that AI algorithms work well in lab- controlled environments, but fail in practical applications in a larger context. Here, suitable measuring instruments must be sought or adapted.
- **Interdisciplinary perspective** regarding the teams that design, develop, deploy, and even use AI applications, thereby combining different subject areas, expertise, and experiences [HMO19].

3 Conclusion

This paper suggests a solution to the currently differentiated development of algorithms under the label of AI in a business environment by suggesting ideas on how to reframe the insights from Software Engineering with the challenges of designing responsible and sustainable artefacts. Some of the perspectives that we put forward, which are certainly not exhaustive, support an interdisciplinary discussion aimed to combine insights from both AI and Software Engineering with a management-oriented view.

References

[Cr21] Crawford, K.: Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press, New Haven, 2021.

[Di19] Dignum, V.: Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Springer, 2019.

[HMO19] Horneman, A.; Mellinger, A.; Ozkaya, I.: AI Engineering: 11 Foundational Practices. Carnegie Mellon University Software Engineering Institute, Pittsburgh, 2019. https://resources.sei.cmu.edu/asset_files/WhitePaper/2019_019_001_634648.pdf

[Jo19] Jordan, M.: Artificial Intelligence—The Revolution Hasn't Happened Yet. Harvard Data Science Review, 1(1), 2019. https://doi.org/10.1162/99608f92.f06c6e61

[La21] Larson, E. J.: The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do. Harvard University Press, Cambridge, MA, 2021.

[MBL20] Musgrave K.; Belongie, S.; Lim, S.: A Metric Learning Reality Check. In Vedaldi, A.; Bischof, H.; Brox, T.; Frahm, J.M. (eds.) Computer Vision – ECCV 2020, Part XXV, Lecture Notes in Computer Science, 12370, Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-58595-2_41

[Mc76] McDermott, D.: Artificial intelligence meets natural stupidity. ACM SIGART Bulletin, 57, pp. 4-9, April 1976. https://doi.org/10.1145/1045339.1045340

[Mc20] McDermott, T.; DeLaurentis, D.; Beling, P.; Blackburn, M.; Bone, M.: AI4SE and SE4AI: A Research Roadmap. InSight Special Feature, 23(1), pp. 8-14, 2020.

[MD19] Marcus, G.; Davis, E.: Rebooting AI: Building Artificial Intelligence We Can Trust. Pantheon Books, New York, 2019.

[Mi19] Mitchell, M.: Artificial Intelligence: A Guide for Thinking Humans. Pelican Random House, UK, 2019.

[MLT20] Monett, D.; Lewis, C. W. P.; Thórisson, K. R. (eds.): Special Issue "On Defining Artificial Intelligence." Journal of Artificial General Intelligence, 11(2), pp. 1-100, 2020.

[Oz20] Ozkaya, I.: What Is Really Different in Engineering AI-Enabled Systems? IEEE Software, pp. 3-6, July/August 2020, DOI: 10.1109/MS.2020.2993662

[Sh56] Shannon, C. E.: The Bandwagon. IRE Transactions on Information Theory, 3, 1956.

[Sm19] Smith, B. C.: The Promise of Artificial Intelligence: Reckoning and Judgment. The MIT Press, Cambridge, MA, 2019.

[So16] Sommerville, I.: Software Engineering, 10th edition. Pearson, 2016.

[Wa19] Wang, P.: On Defining Artificial Intelligence. Journal of Artificial General Intelligence, 10(2), pp. 1-37, 2019.

[Wo20] Wooldridge, M.: The Road to Conscious Machines: The Story of AI. Pelican Random House, UK, 2020.

Author information

Prof. Dr. Dagmar Monets

Sie ist seit 2010 als Professorin für Informatik an der HWR Berlin tätig. Sie studierte Mathematische Kybernetik (Bachelor, 1992), Informatik (Master, 1998) an der Havanna Universität, Kuba, und promovierte in Informatik an der Humboldt-Universität zu Berlin (2005). Im Rahmen ihrer Promotion forschte sie über verteilte KI für die Konfiguration von Metaheuristiken. Sie arbeitete auch am DFG Forschungszentrum MATHEON und am Institut für Mathematik der HU. Aktuell forscht sie aktiv u.a. in den Gebieten KI, digitale Ethik und Informatikausbildung. Sie ist Mitbegründerin von AGISI.org, die Artificial General Intelligence Sentinel Initiative, und akademische Studiengangsleiterin an der Berlin Professional School.

Prof. Dr. Claudia Lemke

Sie ist Professorin für Wirtschaftsinformatik an der HWR Berlin mit mehr als 20 Jahren Lehr- und Forschungserfahrung an verschiedenen nationalen und internationalen Universitäten und Hochschulen. Ihr besonderer Schwerpunkt liegt im Bereich der digitalen Transformation, insbesondere im Kontext der Wirkung digitaler und vernetzter Technologien auf Wirtschaft und Gesellschaft. Zudem beschäftigt sie sich mit der Entwicklung digitaler Lernformate und den Auswirkungen einer digitalisierten Bildung. Gemeinsam mit Dagmar Monet entwickelte sie für den KI-Campus einen Kurs zur Daten- und Algorithmenethik. Claudia Lemke ist akademische Studiengangsleiterin an der Berlin Professional School für einen deutschlandweit einzigartigen Masterstudiengang, dem dualen Master Digitale Transformation.

Workshop Informationen

Workshop Agenda: Management Künstlicher Intelligenz

Montag, 27. September 2021, 10:00 - 18:00 Uhr

Startzeit	Dauer	Thema	Referent
10:00	0:30	Begrüßung und Einführung in das Management der künstlichen Intelligenz	Benjamin van Giffen, Jana Koehler
10:30	0:30	Managerial Cultivation of Training Data for Artificial Intelligence Systems on the Consumer Internet of Things	Jan Zibuschka
11:00	0:30	Pause	
11:30	0:30	Developing IT Systems Based on Machine Learning: Specific Risks and Mitigation Strategies	Benedikt Berger, Johann Kranz; Ioanna Constantiou, Thomas Hess
12:00	0:30	Wie die AUDI AG datenbasierte Technologien für die Serienproduktion entwickelt und erprobt: Das Audi Production Lab	André Sagodi, Henning Löser, Verena Bossdorf
12:30	1:00	Mittagspause	
13:30	0:30	Künstliche Intelligenz für die Smart City	Tabea Hein, Christian Schachtner
14:00	0:45	Diskussion / Zusammenfassung der Ergebnisse früherer Präsentationen	ALL
14:45	0:15	Pause	
15:00	0:30	AI Lab – An Ecosystem for Managing AI projects1	Martin Leucker, Maria Ostanina
15:30	0:30	Towards Reference Architectures for Artificial Intelligence: A Preliminary Report	Alexis T. Bernhard, Jana Koehler, Sophia Saller
16:00	0:30	Pause	
16:30	0:30	AI-ware: Bridging AI and Software Engineering for responsible and sustainable intelligent artefacts	Dagmar Monett, Claudia Lemke
17:00	0:40	Diskussion / Zusammenfassung der Ergebnisse früherer Präsentationen	ALL
17:40	0:20	Zusammenfassung und Abschluss des Workshops	Benjamin van Giffen, Jana Koehler, Walter Brenner, Can Albayrak

Das Organisationsteam

Prof. Dr. Benjamin van Giffen, Universität St.Gallen, (benjamin.vangiffen@unisg.ch)

Prof. Dr. Jana Koehler, DFKI, Universität des Saarlandes, (jana.koehler@dfki.de)

Prof. Dr. Walter Brenner, Universität St.Gallen, (walter.brenner@unisg.ch)

Prof. Dr. Can Adam Albayrak, Hochschule Harz, (calbayrak@hs-harz.de)

Programmkomitee

Prof. Dr. Helmuth Ludwig (Southern Methodist University, Texas)

Prof. Dr. Svenja Falk (Accenture Research/Uni Giessen)

Prof. Dr. Oliver Thomas (Uni Saarland)

Prof. Dr. Dagmar Monett (HWR Berlin)

Prof. Dr. Claudia Lemke (HWR Berlin)

Prof. Dr. Alexander Mädche (KIT Karlsruhe)

Prof. Dr. Nils Urbach (FH Frankfurt)

Prof. Dr. Stefan Morana (Uni Saarland)

Prof. Dr. Thomas Hess (LMU München)

Prof. Dr. Peter Buxmann (TU Darmstadt)

Dr. Benedikt Berger (LMU München)

Dr. Henning Löser (Audi AG)

Dr. Jennifer Hehn (Deutsche Bundesbank)

Management of AI Lab
Institute of Information Management
University of St.Gallen