

International Journal of Electronic Commerce

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/mjec20

Designing a Blockchain-Based Data Market and Pricing Data to Optimize Data Trading and Welfare

Ingrid Bauer-Hänsel, Qianyu Liu, Claudio J. Tessone & Gerhard Schwabe

To cite this article: Ingrid Bauer-Hänsel, Qianyu Liu, Claudio J. Tessone & Gerhard Schwabe (2024) Designing a Blockchain-Based Data Market and Pricing Data to Optimize Data Trading and Welfare, International Journal of Electronic Commerce, 28:1, 3-30, DOI: 10.1080/10864415.2023.2295068

To link to this article: https://doi.org/10.1080/10864415.2023.2295068

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 30 Jan 2024.

Submit your article to this journal

Article views: 816

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=mjec20

INTERNATIONAL JOURNAL OF ELECTRONIC COMMERCE
2024, VOL. 28, NO. 1, 3–30
https://doi.org/10.1080/10864415.2023.2295068

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Designing a Blockchain-Based Data Market and Pricing Data to Optimize Data Trading and Welfare

Ingrid Bauer-Hänsel, Qianyu Liu, Claudio J. Tessone, and Gerhard Schwabe

ABSTRACT

While a wealth of potentially valuable data is generated and stored every year, many businesses suffer from inefficiencies, information asymmetries, and high storage costs, and lack knowledge on how to monetize their data assets. Blockchain is said to offer crucial building blocks to enable a verified, traceable exchange and trading with sensitive data goods and to address current challenges. While the technology's potentials for decentralized data markets have been discussed, the question of how to realize it to optimize trading and welfare remains open. Applying design-science research methods and computational simulation to a real-world business-oriented blockchain project, this study proposes a market model. By adopting the consortium blockchain, we are thinking outside the confines of tokens tied to a blockchain when applying blockchain to the data trading market. Our marketplace is designed outside the speculative tokens space and can focus on the data trading marketplace. We evaluate the effects of different pricing functions on market welfare and trading in on-chain data goods. The results indicate that data trading and welfare can be maximized through a logarithmic pricing function. Further, in a market of heterogeneous agents, unexpectedly, we observe a tipping point in transaction fees above which market operations collapse. Monitoring the market's consumer price elasticity helps us to avoid this collapse node, and we can also impact it by controlling transaction costs. Academics and practitioners can learn about the idiosyncrasies of blockchain in market design and operation.

KEYWORDS AND PHRASES

Blockchain; data market design; distributed ledger; information pricing; data monetization; data trading; data pricing; consortium blockchain

Broader uses of data are expected to generate large economic and social benefits and efficiencies. This is impeded by a lack of data trading and sharing among firms and individuals who fear losing control to centralized data monopolists and unsatisfying compensation for their valuable assets [3, 41, 100]. While some companies' operating platforms already provide data trading services, their target users are mostly limited to large enterprises.

In this context, automotive markets are a particularly interesting and relevant application field, since data in this setting are believed to generate diverse business value for various firms (e.g., insurers, car dealers, or repair shops) [4, 6]. Traditionally, information asymmetries have played important roles in these markets; dishonest deals tend to drive honest deals out of the market [2]. Today, car manufacturers, car repair shops, car dealers, and

CONTACT Ingrid Bauer-Hänsel Ingrid.bauer-haensel@unisg.ch University of St.Gallen, Müller-Friedberg-Strasse 8, St. Gallen 9000, Switzerland

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

insurers already store a wealth of data from a car's history [39], but they cannot mine this treasure because there is no suitable market for it. Even the data sources of Caruso, a platform that focuses on data trading in the auto market, are currently limited to only some of the models of large car companies with onboard equipment [18]. In the car trading market, many tiny sellers (such as individual car owners and tiny repair shops) who possess auto data are also crucial, owing to their huge volumes.

Blockchain technology is said to offer key building blocks to allow for the building of a decentralized market for trusted data exchange [46, 90]. Compared to traditional centralized databases, blockchain both avoids the risks associated with single points of failure and discourages the emergence of data monopolists [28]. It enables one to build standardized infrastructures that allow for distributed access to and control over data [22], while keeping the system control decentralized, which has been deemed crucial for such novel interorganizational collaborations to come alive [98]. Further, it allows for the creation of digital, authenticated tokens (i.e., "a digital, intangible, unique representation of something" [56]). While this something in the Bitcoin system is a digital coin [60], in other systems it could be tokenized data representing, for instance, a property, a mileage update, a repair event, or a hash of externally stored data of a real-world asset such as a car. Such unique representation and management of tokenized data entries could make a market for data attractive [82].

First pilot projects on privately operated business blockchain data markets— particularly in the car ecosystem [7, 63] but also in other domains [37, 38]—have reached a sufficient maturity that makes it worthwhile to ask:

RQ1: How can we design a blockchain-based market for data trading?

To answer this research question, we first need to understand the market's basic logic and structure. We therefore used market design research to gather and analyze information about our target market.

While many important issues must still be solved in blockchain-based data markets—such as security and privacy [46], as well as data quality [35]—one of the most fundamental issues keeping existing projects from moving ahead is financial insecurity. Not knowing and understanding how a business will be able to price its valuable data assets, and therefore how much return on investment these projects can actually bring in, is a showstopper in such projects. Many blockchain consortia [98] and startups struggle to establish and maintain a sustainable income stream. A data market can provide a source of sustainable funding if the money earned at least covers the substantial operational costs. The higher the earnings of any market, the more the offered prices are aligned to data buyers' willingness to pay [89]. In digital markets with high fixed costs and very low variable costs, the pricing of information is particularly challenging [79]. Thus, we separately address pricing as a key decision in the market design, asking:

RQ2: How should trading with on-chain data be priced?

Building on both insights from previous studies and a real-world blockchain project that we collaborated with (hereafter CarData) and that operates a private business blockchain), we applied design science research (DSR) methods and propose a market model for

blockchain-based decentralized data markets. Thus, we focus on a decentralized data market operated by a consortium of businesses that aim to solve business problems in their ecosystem [77]. Given that the project's focus is limited to a specific ecosystem (cars) and that there are strict requirements to comply with data protection regulations, the selected approach is a private consortium blockchain [67], rather than publicly run and operated blockchains such as Ethereum [16, 17] or the Ocean Protocol [64]. We applied computational simulations of a minimalistic agent-based model to analyze how the pricing function influences blockchain-based data trade and welfare. In our model, agents are heterogeneous in their optimal pricing expectations; in these circumstances, it is known that changes in external parameters usually lead to smooth changes in a system's global properties [86]. Our results indicate that the trade of data and the welfare in the blockchain-based data market can be maximized by applying a logarithmic pricing function. However, this also entails the risk of a market collapse if transaction costs are too high (e.g., owing to the high transaction validation costs of a blockchain).

Throughout this article, we use the example of the car ecosystem to more tangibly demonstrate our case; this is also our case project's application domain. Thus, these insights are highly valuable for academics and practitioners in the application realm of the car ecosystem, but are not limited to it. Our contributions can easily be extended to other domains and business models, demonstrating how business-oriented blockchain data markets can be realized. Finally, we provide an answer to calls for research on the design and use of blockchain applications that are informed by and tested in practice [65, 71].

The remainder of this article is structured as follows: First, we review related work on the current problems of and the reason for a lack of data trading and blockchain's potentials and idiosyncrasies for solving these problems. We also review related work on data market design, which our design builds on. We then introduce the applied methods and data collection processes. Subsequently, we present our results, including the objectives and design requirements we derived from previous work and the CarData project for the market model; our final market model design; and the simulation results, which show the effects of different price functions on the trading and welfare of on-chain data. Finally, we discuss the results and conclude by highlighting the study's contributions, limitations, and future research avenues.

Background and Related Work

In this section, using the car ecosystem as an example, we dissect the need for and current dilemmas in data trading, and discuss the solutions proposed by relevant research, as well as the possible problems with them.

Predicament

Despite the almost ubiquitous generation and potential availability of rich and valuable data sets, they are often not shared or traded [41, 100]. Thus, many firms and private users suffer from inefficient processes [10], information asymmetry [63], and large storage capacity requirements without the actual data being utilized [57].

In the car ecosystem, for instance, when a car is imported, the import form changes hands multiple times between the customers and car businesses (e.g., importers, road-traffic

authorities, insurers) and also between these businesses themselves. Such processes are cumbersome for customers in terms of the time and work they have to invest to provide their data to multiple businesses multiple times. They are also inefficient, prone to error, and rigid from a business perspective. Each of these parties collects similar basic data sets about a car and its owner and accrues the data relevant to its internal processes, leading to double (or triple, or other multiples) work in data recording and multiple susceptibilities to errors [5]. Information asymmetries provide another significant challenge in the car ecosystem, but also in other industries [2, 4, 85]. Insurers, for instance, lack insights about the customer types who request insurance; also, they often do not know the value of the car that they are insuring [9]. Thus, they struggle to set fair insurance premiums that reflect a car's value yet do not lead to adverse customer selection [2]. Further, private buyers and sellers of used cars struggle with information asymmetries and their negative consequences during the sale of a used car [12]. As described in reference [2], in markets with an asymmetric distribution of information, the party who has more information—in this case, used-car sellers—levers their information advantage and trades cars of low quality for prices that do not justify their quality [14]. This leads to negative consequences of adverse selection. Another issue is that businesses with large valuable data sets struggle with storage and do not know how to monetize their resources [41].

While there is strong interest in exchanging and trading data, many practical dilemmas remain. On the one hand, from a data buyer perspective, data sets are unstructured and scattered, while the quality and their origin are unclear [52, 97]. Further, data buyers are interested not only in data provided by one seller but also in the combination of data sets from multiple data providers [57, 96]. Finally, businesses that seek to utilize the data for regulated processes (e.g., document handling) need to be sure about data's trustworthiness and traceability [59, 66]. On the other hand, from the data seller perspective, the reason for few data trading incentives arises from the challenge to price data [87]. Since data are considered an easily replicable good with high initial production costs and marginal replication costs, it is hard to find an adequate pricing mechanism that is not prone to arbitrage opportunities [79,87]. This is an even bigger problem in cases such as the car ecosystem if the buyers' values are not known and if buyers' values vary for different combinations of data sets [41]. While the cases just mentioned cases (process inefficiencies and information asymmetries) provide some hints for data sellers about buyers' values, the ability to utilize the data insights for other, as-yet-unknown cases makes it very difficult for data sellers to estimate buyers' true values. Thus, many data sellers are hesitant about getting involved in data trading today.

Solutions and Limitations

In recent years, many platforms that offer data trading services have emerged around the world. Their data trading models can be classified into three broad categories: unilateral data provision model, data trading platform model, and data management system model [3]. Most of these platforms are still centralized data marketplaces; that is, the platforms are operated by intermediaries that undertake the collection, processing, and resale of data [76]. Prominent examples of such many-to-many data markets operated by intermediaries [41] in the car ecosystem are Carfax and Eurotax. While these centralized data markets resolve buyers' issues to a degree (i.e., aggregating and structuring data from multiple sources and

providing a quality label), there are limitations in firm-controlled marketplaces, from a seller perspective: The firms are more focused on maximizing their interests, neglecting the protection of and incentives for sellers [81]. Besides the already-mentioned arbitrage opportunities and thus few incentives for sellers to provide and sell their data to such intermediaries, there is the fear of establishing a novel centralized power, such as Google or Facebook [98]. These fears are supported by regulatory pressures (e.g., by the European Union) that call for the dismantling of platform giants [75] and the building of a government-backed data trading marketplace without platform giants. According to the World Economic Forum, India, Colombia, Japan, and China are experimenting with data marketplaces, but the concept is still very nascent [83, 92]. Rather than having monopolistic intermediary data giants, data sellers may want to retain control over their data (usage). Further, buyers lack trust in the quality of the data provided by single intermediaries [97]. Thus, the call to adopt decentralized data markets is growing louder, challenging established platform intermediaries [41, 76, 102].

Federated learning or federated computing may be a good solution for the decentralized data trading market. It was proposed by Google in 2016 and allows efficient machine learning to unfold across multiple computing nodes while ensuring user privacy [53] and security. Each node can benefit from this data sharing [54]. However, federated learning still faces many problems in application. On the one hand, the existing network bandwidth between user terminals and enterprises does not meet federated learning's needs, and the technical threshold is high [54]. It is not conducive to involve large-scale tiny participants. In the car ecosystem, even the vehicle data of individuals and historical data of a large number of tiny garages and used-car dealers are in high demand in the market. Meanwhile, federated learning suffers from the same problems of single-point failure and uncontrollable data quality as centralized data markets [47]. Many studies have tried to overcome this problem and have proposed combining federated learning with blockchain technology to guarantee stable data sharing through more decentralized data and computing power, and expanding the market size and guaranteeing the market dynamics through incentive mechanisms [62]. The authors of reference [80] have shown that, in the context of CarData, it is possible to incentivize (through rewards tied to a *reputation* system) the participants' data quality and minimize deviant behaviors.

Today, blockchain has been proposed as a technology with the potential to resolve the already-mentioned challenges and provide new opportunities through decentralized, verified (sometimes referred to as trusted) data trading [7, 21, 90]. Technically speaking, blockchain builds on a distributed database architecture that validates and stores transactions in a decentralized way without the need for a central, trusted authority [10, 24]. Participants in the network reach consensus along predefined rules, and the technology's inherent cryptographic logic defines its unprecedented tamper resistance [44, 101].

While advocates derive multiple common benefits from such a technical view, concerning a data market, two key advantages can be identified: First, as the foundational technology for a reliable and distributed record system, it enables new types of interpersonal and interorganizational relationships in a distributed, decentralized way [23, 78]. Second, blockchains allow one to create a form of unique ownership and traceability through trading in digital goods such as data [55, 56, 82, 94], increasing trust in digital assets [51, 97]. Blockchain can create unique control and exchange of data records (e.g., regular mileage updates that provide details about a car's history), the immutability and proof-of-

provenance of which have motivated many businesses to explore and build novel business-oriented blockchain-based data markets [20, 90, 98]. Examples include the health sector [31, 99], the car ecosystem [7, 63], the shipping sector [38, 59, 74], and, in short, any sector that seeks to digitally manage valuable assets [70].

Blockchain technology has the potential to establish a truly reliable market and to uniquely manage digital assets, establishing a sustainable business model, which is hard yet necessary for successful adoption [43, 95]. A case study of blockchain use in regulatory fields [30] has demonstrated the potentials of using the technology for regulatory compliance, with the data quality management process as a good reference use case. Since we focus on the trading of car data, which does not need to be open to everyone, we consider using a consortium blockchain to build a data marketplace to effectively reduce operational and trust costs. The major challenge for business-oriented operators of blockchain data markets is the high operational costs of data storage and eventual data processing [25]. While these costs are not as high as in a permissionless blockchain, data-centric operations are not cheap. In contrast to centralized databases, the distributed verification model inherently demands more transaction fees, which may endanger some business models. Thus, how a decentralized model will be worth implementing is a key question for business. Not knowing whether investments will provide adequate returns is often a key hindrance in such projects, besides privacy and security issues [46]. Also, the revenue achieved in such decentralized markets needs to be distributed fairly among different data providers and system operators if one is to create and maintain a self-contained system [29]. Further, owing to its multifaceted nature, both data sellers and data buyers are needed simultaneously to achieve a critical mass and then a high system closedness [24,29]. Thus, maximizing welfare for the entire ecosystem rather than for one user group is an objective that is also required by its decentrality [11], especially among business-oriented blockchain consortia. The research has largely focused on the specifics of its technical implementation and proofs of concept [71]. However, to deploy and establish such data markets in practice, the vital next step is to understand and propose a proper market design and evaluate its feasibility for certain business models. To our best knowledge, this has not yet been addressed.

We focus on the design of a blockchain-based data trading market, without discussing too many details related to blockchain technology. Based on this discussion, the market design has three primary assumptions, owing to the consortium blockchain's benefits.

First, the market's management consortium establishes the blockchain. The on-chain traders are permissioned, there is no significant cost (proof-of work etc.) to maintain trust within the system, and the data pricing only considers the transactions. Second, the data in the transaction are traceable, ownership of the data can be determined, and the buyer only has the right to use the data without ownership and cannot resell them for arbitrage. Third, we specifically use direct fiat pricing rather than tokenized pricing in the context of the consortium blockchain. This can greatly reduce the concerns of excessive cost, energy consumption, and speculation when applying blockchain. Table 1 summarizes the comparison of the different solutions.

Data Market Design

Prior research on data markets provides valuable insights for solution design. However, they are either too general to apply to specific instantiations of the described problems or too specific, and the results cannot be extended to other cases with different requirements

Table 1. Comparison of Different Solutions

	Traditional data trading platforms	Federated learning	Consortium blockchain
Centralization	Yes	A bit	No
Monopoly risk	High	Low	Low
Single point of failure risk	High	High	Low
Technical thresholds for traders	Low	High	Low
Privacy	Low	High	High
Data quality	Trustless	Trustless	Controllable
Traceability	No	No	Yes
Data validation	Lack	Lack	Yes
Proof of provenance	Lack	Lack	Yes
Transaction cost	Low	Low	Controllable
Trust cost	High	High	Low
Remarks		Can be combined with blockchain	

and objectives. To our best knowledge, reference [57] is the closest work regarding the requirements described in the related work and the requirements we derived from CarData and specify in the subsequent sections. Its authors propose a design of a combinatorial data market, where data buyers can request combinations of data sets from multiple buyers. Further, they do not limit the buyers' data requests to a certain application and design for combinatorial requests where the buyers' values are as yet unknown. Further, they consider data as a replicable good, while blockchain data markets may provide unique ownership to data and thus create a type of scarcity that is generally known to impact on consumer demand [36]. While data can in principle be copied and resold outside the blockchain system, they will not hold the same value for users. Also, the authors design the market to be buyer-optimal, arguing that the lack of buyers is a major reason why combinatorial data markets—such as the Azure Marketplace owned by Microsoft—have failed [57].

Another recent data market design comes from reference [1], which designed a two-sided market for prediction task data sets. The authors design mechanisms for a combinatorial data set and provide valuable insights into mechanisms designed for fair revenue distribution. They propose a mechanism that incentivizes buyers to truthfully report the value of data. Specifically, buyers are asked to report a bid for the data sets they are interested in, and the payment is then computed using methods from Myerson [104] and a zero-regret objective. While reference [1] could design a truthful mechanism, its mechanism lowers the quality of the data sets sold in case of bids below the buyer's real value. That is, concept blockchain-based data markets cannot adapt, again given the tokenized data sets' uniqueness [19, 56]. Further, the truthful mechanism and the extracted buyer's value let the authors disconnect the buyer's utility from the purchased data sets. The price calculation is merely based on the buyer's value, which in turn depends on the buyer's application of the data set. However, as noted, in blockchain-based data markets, buyers' interest in such data markets lies in the diversity of their application potential [6]. Hence, limiting the buyer in their use of a data set to certain applications is not an option in a general-purpose approach in blockchain-based data markets.

A different approach is proposed by the authors of reference [40], who do not set any restrictions on the buyer or seller side. Further, they consider data as a nonreplicable good, and therefore arbitrage-free. However, in this setting, sellers are required to price the data themselves. This may be a reasonable approach in some settings; however, in the use cases

considered in this contribution, data sellers are believed to lack the necessary knowledge to do so. Thus, surrendering the difficulty of data pricing to sellers might end fatally for the whole ecosystem.

Federated learning, as a decentralized solution for data trading markets, also provides a good reference for data value determination based on it. The federated Shapley value [91] is a measure of data value in the federated learning framework that satisfies many desirable properties for data valuation. Building on this, a new measure—the completed federated Shapley value—was introduced in reference [26], resolving potential unfairness in the design of federated Shapley values very well, both theoretically and empirically, by introducing a novel utility matrix concept. However, in our case, there were no restrictions on the buyer's use of the data set, especially for small and micro sellers and buyers, who may need to directly obtain or sell clear data from a small number of target users in order to optimize decision making. The value difference between individual data points is very small, and the volume of data may strongly impact on the price.

Finally, an important question that pervades market design is how to price exchanged goods when an objective—such as welfare maximization—is predetermined. A survey on data pricing [69] provides a unified, interdisciplinary, and comprehensive overview over data pricing, summarizing the basic principles of data pricing: versioning, truthfulness, revenue maximization, fairness, no-arbitrage pricing, and privacy protection. As discussed, some of these data pricing principles are naturally satisfied in the consortium blockchain context. Thus, while we only need to think about revenue maximization, we want to go one step further and think about maximizing benefits.

The task difficulty even increases if the consumers' demand functions are not only different but very heterogeneous, as is the case in the automotive ecosystem. Further, in the specific case of a blockchain data market, each data provider has a monopoly on the data it controls in the platform [90]. In these conditions, the simplest possible approach is to apply average cost pricing, which—while allowing for keeping space for future investments—leaves consumption at a suboptimal level [15]. A simple extension of this approach is a two-part tariff [27], which is valid when the average cost is a decreasing function of the demand. Here, prices lower than the average cost may be applied, while the balance is recovered through a basic license fee, irrespective of the consumption. While in many examples the license fee was set to zero [8,50], notably, reference [61] showed that the optimal price "may be below marginal cost if consumers at the margin have above average consumption," for instance, in highly heterogeneous environments where many small consumers with very steep demand curves keep consumption low, while those with large consumption have a flat demand curve.

Despite certain shortcomings, these studies provide valuable foundational work; our design builds on them.

Methods

To design and evaluate a market model for blockchain-based data trade, we applied design science research methods [33]. We also used an agent-based modeling (ABM) approach to simulate the actions and interactions of autonomous agents (buyers and providers) in data trading markets so as to understand the behavior of the system and the factors that control its outcomes [48].

DSR originates from the computer and engineering sciences and seeks to solve practical problems through the design of novel artifacts [49, 68]. These artifacts can take the form of concepts, models, or instantiations [33]. Our study was initially motivated by the problems we observed in the CarData project, which we collaborated with for more than 2 years, attending several project workshops, including steering committee meetings, project status meetings, and design and development meetings with the scrum development team. Cardossier is a real-world project initiated by a consortium of firms from the car ecosystem in Switzerland. This includes car importers, road-traffic authorities, insurers, mobility services, and (car) data analytics firms. The project's goal is to address various issues in the car ecosystem (e.g., inefficiencies, information asymmetries, unused resources) by building a decentralized, blockchain-based data market. Yet the project lacks insights about how the market mechanisms should be designed and how data on blockchain should be priced. Applying DSR methods, we conducted two design and evaluation iterations [33], as described in the following and as summarized in Table 2. Our results include a validated understanding of the problem and a specification of the design requirements and market objectives, the design of a market model for data trading, and market simulations that allowed us to test different pricing functions. We will report our findings and will add them to the prior knowledge base [32].

Iteration 1: Motivated by the problems we observed through our collaboration with CarData, we abstracted from the instance problem [42] and studied the literature on abstract problems and the reasons for the current general lack of data trading. Further, we studied the literature on blockchain to understand its idiosyncrasies for market design and data trading. To verify and refine our understanding of the problem and to derive specific design requirements and market objectives that a market model would need to meet, we again de-abstracted the instance case and conducted interviews with several CarData project members. This review of the literature and the project formed the basis for understanding the problem and the requirements and market objectives that our design had to meet, which led to Iteration 1's results.

Iteration 2: Based on this thorough understanding of the problem space, we reviewed the literature on data market design. Building on previous work, we then designed the Permissioned Blockchain Data Market (hereafter PeBDaMa) model, which considers the requirements and market objectives identified during Iteration 1. Further, we implemented the market model and ran computer simulations to study how the trading and welfare of data can be maximized [33, 88]. The actions of the agents in our setup were rigidly fixed by the conditions they were exposed to. Thus, the approach can also be understood as a simplified agent-based model [13, 93]. Finally, we again conducted interviews with firms in the CarData consortium to evaluate the PeBDaMa model and discussed the simulation results.

We followed a semistructured interview approach [58], and recorded and analyzed the interviews. Two researchers developed the interview guides and conducted the interviews. The interviews lasted between 35 and 50 minutes; they were recorded and documented, and the key findings were summarized and sent to the participants for validation. After they confirmed the documented interview transcripts, we again analyzed the interviews, in pairs of two. Two researchers first applied an individual initial open-coding process, then jointly discussed their findings [73].

Table 2. Overview of the Evaluation Phases and the Participants

Phase	Meeting objective	Firm		Participant(s)	Date
Evaluation 1	Validation of the understanding of the problem and specification of design requirements and market objectives	A	Car importer, dealer, and repairer	1. Director of Business Transformation	May 2019
		B	Insurer	1. Head of Innovation Engineering 2. Internal Consultant	May 2019
		C	Car data analytics firm	1. CEO	July 2019
Evaluation 2	Evaluation of the final market model and the simulations' results	D	Data analytics and IT firm	1. CEO 2. Lead Product Manager 3. Lead IT Architect	August 2019
		A	Car importer, dealer, and repairer	1. Director of Business Transformation	September 2019
		B	Insurer	1. Head of Innovation Engineering 2. Internal Consultant	September 2019
		C	Car data analytics firm	1. CEO	October 2019

We then revisited our model and the simulations, performing robustness checks, making improvements to prior fixed variables, and rerunning the simulations. The conclusions of our analysis remained unchanged. (Documentation of these steps is available upon request.)

Results

The CarData Project: Objectives and Design Requirements

In the following, we report on the design requirements and objectives we derived from the work and from our interviews with CarData project members.

Design Requirements

The firms in the CarData project confirmed that multiple processes that comprise cross-organizational document handling suffer from costly inefficiencies that rapidly increase with the number of stakeholders involved. Interviewee A.1 mentioned that importers, for instance, possess abundant amounts of valuable data about cars they had imported (e.g., original price, installation parts); however, they struggle to price their data assets. Also, the partners confirmed that the repair history or information about past driving dynamics is valuable during the sale of used cars, yet data requests from buyers and the value they accorded to them differed greatly. In sum, the project partners emphasized that potential combinations of data requests from buyers were abundant, and that the buyers' values were unknown and were expected to vary for different combinations. Thus, for buyers' values, we derive:

Requirement 1 (buyers' values): Data needs to be priced below unknown buyers' values, which vary for different combinations of data requests.

While the exact buyers' values for data requests are yet unknown, from the interviews we derived that the stakeholders expect a Pareto distribution of transactions in the market. For instance, Interviewee B.2 mentioned that *"gaining access to these data sets could enable to resolve information asymmetries for individuals but also issues that stem from a lack of customer insights as well as inefficiencies for businesses."* In their view, this will generate many small individual requests from private used-car buyers and sellers, as well as a few large bulk requests from larger cooperates. Adding to this, Interviewee C.1 stated that in their business they are already seeing a Pareto distribution in demand: *"Assuming that private buyers participate once every few years, they will have fewer than ten transactions, inducing rather few data records specific to a car. However, these private buyers will be the majority of buyers in the market. On the other side of the spectrum are large commercial buyers . . . who will be interested in a few hundreds or thousands of data records per year in rather few bulk requests. In between, we expect a middle ground of other smaller institutional buyers that complete several transactions per year."* Accordingly, we derive requirement 2:

Requirement 2 (buyers' demand): Buyers' transaction demand represents a Pareto distribution.

Besides the challenges of businesses in the car ecosystem, CarData also struggles with the question how its development and maintenance costs will be covered. As described in the literature, also from the project meetings, we derived that the market must be able

to finance itself. Currently, CarData relies on membership fees to cover the development costs. However, as mentioned by Interviewee C.1, in the future the market should be self-sustaining.

Requirement 3 (platform business model): The market should be self-sustaining.

Market Objectives

The project analysis revealed that, especially in the beginning, CarData sought to achieve a critical mass of users as swiftly as possible. As Interviewee A.1 stated, "The team should aim to achieve as much traffic in the market as possible, to foster the completeness of all car histories." Thus, first and foremost, maximizing trade was confirmed as the main market objective (C.1), followed by the aim to maximize total welfare in the entire ecosystem (B.2, A.1). Thus, we derive the market objectives:

Objective 1 (data trade): First and foremost, the market mechanism should maximize data trading.

Objective 2 (total welfare): The market mechanisms should maximize total welfare in the market.

A Consortium Blockchain-based Market Model for Trading Car Data

The proposed market model for car data trading (PeBDaMa), which evolved from the insights from CarData and from previous studies (presented in the *Background and Related Work* section), collects and maintains car data from multiple stakeholders. The basic system structure is built on a blockchain-based ledger and allows agents to request data in various combinations (R1). We next explain the key functions of the data market and the optimization objectives that guide its design.

Agents

PeBDaMa contains a set R of records. As the underlying blockchain infrastructure follows an append-only structure, the value of N_R increases over time, as new blocks are added to the blockchain.

In the market, we differentiate between *two agent types*:

- *Buyers* are agents who request data records from the market. At time t buyer b can request any combination of data records published in the blockchain $S^p(t)$, and assigns an individual value $v_b\ (t) \geq\ \mathbb{R}^+$to them (R1).
- *Providers* are agents who deliver data records by publishing data into the blockchain. Examples include car-related businesses or data registrars. A provider m has a cost $c_m\ (t) \geq\ \mathbb{R}^+$ for delivering the data contained in transaction t. These costs may represent registration, infrastructure maintenance, and so on. In return, providers receive revenue for delivering data to the market. The data may (or may not) belong to the provider.

While we differentiate between these agent types, an individual agent can fulfill any combination of these roles. PeBDaMa is the meeting ground for the agents and controls the exchange of data and money. To be self-sustainable [29] (R4), the market will charge

a fee for each transaction, as is normal in blockchain-based systems [84]. This fee covers the total operational cost o_c incurred by the market for its operation. This additional ingredient allowed us to investigate market costs' role in a market's welfare.

Interactions Between Agents

A transaction starts its journey when a buyer arrives in the data market and shows interest in buying a set of data records that have been published in the blockchain[1] $S(t) \subseteq R$ (R2). Sellers can only sell published records. Nonpublished records cannot be bought. Thus, |S(t)| denotes the number of records requested in a single transaction. We assume that the distribution of the number of records per transaction $|\mathrm{S(t)}| \tilde{\mathrm{Pareto}}(\gamma_r)$; that is, it follows a skewed, broad distribution, reflecting the diverse nature of the buyers. This assumption is parsimonious, as buyers may require records about a single or many cars, depending on their use. Every buyer b can execute an arbitrary number of transactions. We denote the total number of transactions by all buyers as N_T.

To compute the *buyers' value for a transaction* $v_b(t)$, we build on the requirement (R1): On the one hand, as buyers' values are unknown and we assume all records have the same value, we consider that the cost that a buyer b is willing to pay for the set of records $S(t)$ follows a monotonic increasing function on the number of items, which is the same for all buyers. Specifically, we consider the following profiles: (a) a logarithmic value function $\alpha_b \log(|S(t)|)$; (b) a linear value function $\alpha_b |S(t)|$; and (c) a constant buyer value α_b. Nonetheless, we assume that α_b is an idiosyncratic term that is drawn from a distribution for each buyer b, characterizing the whole population. A robust property of economic actors is that, independent of different scales (e.g., individual disposable income, firm revenue distribution, etc.), all follow a Pareto distribution [72]. Thus, we assume that all buyers have the capacity to spend in PeBDaMa following a Pareto distribution with exponent η_b, that is, $B_b \tilde{\mathrm{Pareto}}(\gamma_b)$. We make use of this universal feature to perform a parsimonious modeling assumption. This allows us to write that the rate at which buyers request transaction records is proportional to B_b. We assume that α_b is equal to a fraction of the value of the car (or average over all the cars in a transaction); thus, for each buyer $\alpha_b \tilde{\mathrm{Normal}}(\alpha_{mean}, \alpha_{std})$ is distributed according to a normal distribution with given mean and standard deviation. Also, we introduce a buyer price tolerance δ, defined as the fraction of overpricing the agents are willing to pay with respect to what they consider as a fair price, as per their pricing function and idiosyncratic parameters.

For *pricing transactions* (i.e., the price set by the market for a request from buyers for data records) in the market, we build on prior introduced mechanisms that are the closest to our use cases.

Posted price only: As in the setting of reference [57], a buyer pays an issued price for a single transaction. The price is a direct function of the number of published data records, denoted as $|S(t)|$, that the buyer is asking for in their transaction t. Further, given unknown buyers' values and countless possible data requests [6, 57] (R1), we have to identify a suitable pricing function. Therefore, we examine three pricing function types, where we assume that price is a monotonic function of the number of $|S(t)|$.

[1]For simplicity, and without loss of generality, we consider that buyers are only interested in published records.

Constant price function $pr_c : \mathbb{N} \rightarrow \mathbb{R}$. The price per transaction does not change with the number of records a transaction contains. Thus, $pr_c(|\ S(t)|) = \beta_c$, where β_c denotes the price that a buyer must pay for each transaction regardless of the amount of data it contains.

Linear price function $pr_l : \mathbb{N} \rightarrow \mathbb{R}$. The price changes linearly with the number of published and valuable records bought. The function $pr_l(|\ S(t)|) = \alpha_l|\ S(t)| + \beta_l$ has two parameters. First, α_l denotes the (linear) marginal cost of each additional record. Second, β_l denotes a minimum price a buyer must pay to post a query (even if devoid of data).

Logarithmic price function $pr_a : \mathbb{N} \rightarrow \mathbb{R}$. In this case, each additional record has a decreasing marginal cost on the number of published records bought. We again have two parameters in the function $pr_a(|S(t)|) = \alpha_a \log(|S(t)|) + \beta_a$ that influence the price of the records when appending an additional record, and β_a that serves as an initial threshold as it does for the linear pricing function.

In PeBDaMa, the decision by a buyer b to complete (or not) a transaction depends on whether they are below or above $v_b(t)$ the price $pr(t)$ of the transaction. We write the function $ct(t)$ as

$$ct(t) = \begin{cases} 1 & pr(t) \leq (1+\delta\ )\ v_b(t) \\ 0 & pr(t) > (1+\delta\ )\ v_b(t) \end{cases},$$

determining whether or not the transaction was completed. If completed, a buyer b must pay the total price of $PR_b = \sum_{t/\beta(t)=b}\ pr(t)\ ct(t)$ for all transactions they performed.

Global Properties

The main *market objective* is to maximize car data trade (O1) on the market and subsequently to maximize the total welfare through car data traded in the ecosystem (O2). Formally, we define the market objectives as follows:

Total trade: The total number of completed transactions is given by

$$\mathrm{T}_T = \sum\nolimits_t\ ct(t)\,,$$

while the total number of exchanged records is given by

$$\mathrm{T}_R = \sum\nolimits_t\ |\mathrm{S}(t)|ct(t)$$

As per the market requirements, the number of transactions completed and the records exchanged should be maximized (O1).

Total welfare: Given the data market, the total extent of welfare among all agents should be maximized. The total surplus in a market is a measure of the total welfare of all participants in it; it is the sum of consumer surplus and producer surplus [34]. Consumer surplus is the difference between the price a consumer is willing to pay for a good and the price the consumer pays. The producer's surplus represents the difference between the amount received by the seller and the price they are willing to sell each unit of an item for. In this system, buyers are on the consumer side, and their surplus (welfare) is the price the buyer is willing to pay for a transaction minus the price paid for completed transactions:

$$\mathrm{W}_b = \sum\nolimits_t\ (v_b(t) - pr(\mathrm{t}))\ ct(t)\,.$$

Table 3. Overview of the Parameters

Parameter	Description
R	A set of data records.
N_R	Total number of records in the market. In the simulations, we fixed $N_R = 10,000$.
b	Buyer agents can request data records from the market. In the simulations, we fixed the number of buyers at 500.
m	Provider agents who deliver data records to the market. We fixed the number of providers at 500 in the simulation.
$v_b(t)$	The individual value that a buyer assigns to a set of data records.
$c_m(t)$	The costs that a provider has for delivering the data contained in transaction t to the market.
N_T	Total number of transactions requested by the buyers in the market. In the simulations, we fixed $N_T = 50,000$.
$S^p(t)$	Set of published records in transaction t.
pr_c	Constant price function.
pr_l	Linear price function.
pr_a	Logarithmic price function.
o_c	Total operational costs of the market.
$ct(t)$	Determines whether or not a transaction was completed.
η_b	Exponent of the Pareto distribution of the buyers' spending capacity in the PeBDaMa. In the simulation, η_b is fixed at 2.
a_b	Parameter drawn from a distribution for each buyer b, characterizing the population as a whole.
a_m	Linear parameter used by the market for pricing transactions that best fit buyers' expectations.
β_m	The amount a user must pay, even if it contains no records.
δ	The overpriced that agents are willing to pay in relation to what they consider a fair price. In the simulations, we fixed $\delta = 10\%$.

Providers are on the producer side; owing to the special properties of data in the car industry, the cost of transacting data on the chain can be simplified to transaction fees only, regardless of the cost of the data. Thus, their total welfare is the total revenue generated by them:

$$\mathrm{W}_p = \sum\nolimits_t \left(pr(\mathrm{t}) - \beta_m\right)\ ct(t)\,.$$

This means that the total net welfare is

$$\mathrm{W}_T = \ \mathrm{W}_p + \ \mathrm{W}_b$$

The total revenue for the platform is simply given by $\sum_t \beta_m\ ct(t)$.

An overview over all parameters just introduced is shown in Table 3.

Simulation of the Market Dynamics

We simulate the model in three phases: agent setup, market setup, and market operation.

Phase 1: Agent Setup

Given the number of buyers and providers, and the parameters that determine them, they are initialized; that is, each buyer is endowed with the idiosyncratic parameters that determine their fair value for transactions. The pricing function for the consortium who governs the market is also a parameter.

Phase 2: Market Setup

The consortium governing the market sets it up to allow for fluid operation. They have no prior information about the pricing function used by the buyers, their activity distribution, or that of idiosyncratic parameters. To set up the market, N_T transactions with $|\mathrm{S(t)}|$ records

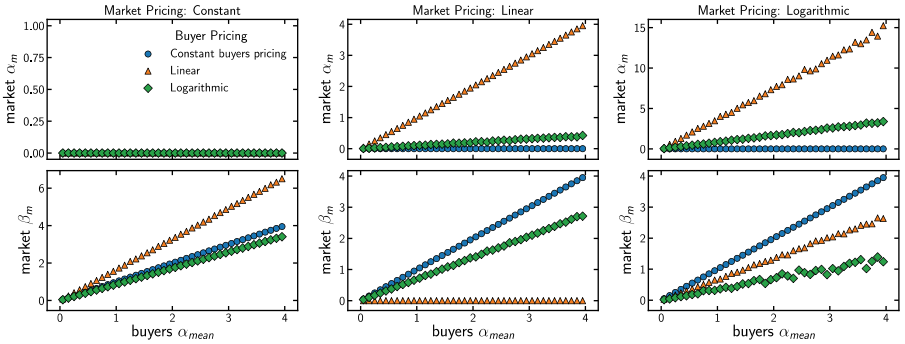


Figure 1. The Market's Parameters, as Regressed as a Function of the Mean of the Buyers, With Market Parameters a_{m} (Upper Row) and β_{m} (Lower Row) as Regressed as a Function of the Mean $a_{\mathrm{b,mean}}$ of the Buyers, Where Different Curves Represent Different Pricing Functions by the Buyers, While Different Columns Depict Results for Constant, Linear, and Logarithmic Market Pricing (from l to *r*)

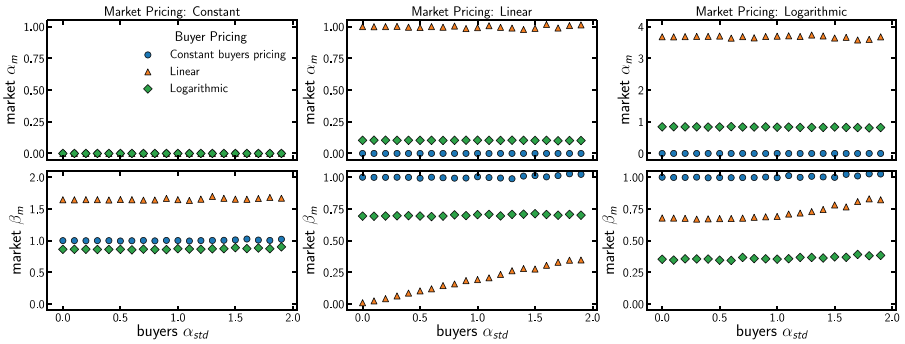


Figure 2. The Market's Parameters, as Regressed as a Function of the Standard Deviation of the Buyers, With Market Parameters a_{m} (Upper Row) and β_{m} (Lower Row) as Regressed as a Function of the ç $a_{\mathrm{b,std}}$ of the Buyers, Where Different Curves Represent Different Pricing Functions by the Buyers, and Different Columns Depict the Results for Constant, Linear, and Logarithmic Market Pricing (from l to *r*)

(sampled from the corresponding Pareto distribution) are generated. Each transaction is assigned to a randomly selected buyer (with a likelihood proportional to their activity). Every transaction's price is queried by the buyers. This generates a sample or representative transaction sizes and the prices buyers are willing to pay for them.

The market operator selects a pricing function (independent of the one used by the buyers) and estimates the values of α_m and β_m^0 that best fit the sampled data. Given the operational cost o_c, the market operators set $\beta_m = \max\left(\beta_m^0, \frac{o_c}{N_T}\right)$, to cover the costs. Thus, the value assigned by the market to each transaction is generated endogenously.

Phase 3: Market Operation

The market then runs for N_T transactions. For every transaction, a buyer is sampled with a probability proportional to its activity. The number of records in this transaction is

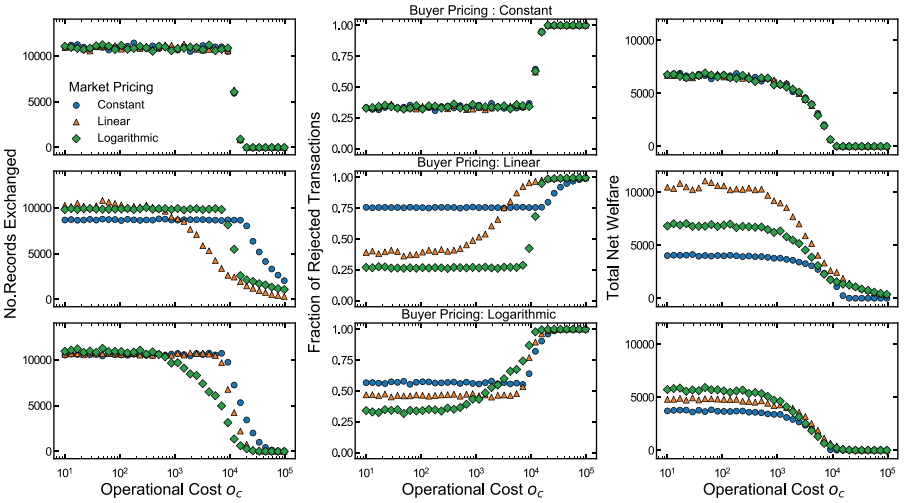


Figure 3. Global Properties of the Market as a Function of the Operational Cost, cost_c, Where the Upper, Middle, and Lower Panels Show the Results, Respectively, for Constant, Linear, and Logarithmic Pricing Functions, Followed by the Buyers, and the Left-Hand Panels Show the Number of Records Exchanged, the Middle Column the Number of Rejected Transactions, and the Right-Hand Column the Net Welfare, wlth Different Curves Representing Scenarios where The Market Implements Constant, Linear, or Logarithmic Pricing Functions as a Function of the Number of Requested Records

sampled from the corresponding Pareto distribution. The buyer queries the market for the transaction price. The transaction is completed based on the conditions explained in the objectives and design requirements. All measures are updated, and the process starts again.

Model Calibration

To instantiate the market and examine the effects of different pricing functions, we first simulate the determination of α_m and β_m by the market, provided a set of buyers (as explained in the Phase 1 of the simulation of the market dynamics). After N_T transactions, the market selects a pricing function for itself and estimates the parameters α and β. The results are shown in Figure 1 for α and in Figure 2 for β. Notably, there is a bias in the sample owing to the highly skewed distribution of the utilized buyer activity. However, we see that for different values of $\alpha_{b,mean}$ and $\alpha_{b,std}$, the results were fairly stable. The estimated values of α_m and β_m depend roughly linearly on the changes in the underlying distributions of buyers' idiosyncratic terms. Thus, in the following, we fix $\alpha_{b,\ mean} = 1$, $\alpha_{b,std} = 0.25$. We also fix the buyer price tolerance $\delta = 10\%$, without any loss of generality (the results were similar for a broad range of values). (Further details about the model instantiation are available upon request.)

Effects of Pricing Functions on Data Trade and Welfare

Taking the market operational costs o_c as the independent variable, we now investigate the different pricing functions' roles in the overall market performance. The market's sole purpose is to provide an environment for data providers and data buyers to meet in. After

estimating the parameters that best match the buyers' preferences, the market may need to set a higher transaction fee if this is required to sustain its operations. The results are shown in Figure 3.

Specifically, in these plots, we show different global properties of the market as a function of the operational cost o_c. The upper, central, and lower panels show the results for constant, linear, and logarithmic pricing functions followed by the buyers. The left-hand panels show the number of records exchanged, the central column the number of rejected transactions, and the right-hand column the total net market welfare. The different curves in each panel represent scenarios where the market applies constant, linear, and logarithmic pricing functions as a function of the number of records requested. We observe the following:

First—for the parameters selected—when $o_c > 20{,}000$, the market needs to charge a larger transaction fee than the one acceptable to the buyers, irrespective of their (and the market's) pricing function. We see that the fraction of rejected transactions is larger when the market applies a constant pricing function (which performs very poorly for linear buyers' pricing function). A market's logarithmic pricing function produces (for low market running costs) the lowest number of rejected transactions. Concerning the number of records exchanged, we see that for very low running costs there is a regime in which linear and logarithmic market prices attain similar numbers of transacted records.

Regarding the market balance, once again applying constant pricing produces larger losses, which are minimized for logarithmic pricing functions for operational costs lower than 10,000. The buyer expenditure is roughly constant for the same range of o_c. However, we also see that the logarithmic pricing function causes higher costs for logarithmically priced buyers, and lower costs for linearly priced buyers. Then (for a logarithmically priced market), the total net welfare is larger for logarithmic priced buyers, and equal for constant and lower for linearly priced buyers.

So far, we have concentrated on the regime of fairly low operational costs ($o_c < 10{,}000$). Interestingly, when the market needs to apply a higher transaction fee, the number of accepted transactions collapses for most combinations of parameters; the buyers are therefore inelastic to changes in the fee structure. Notably, even slight variations can induce a sudden increase in rejected transactions, as well as a collapse of total net welfare and the number of records exchanged. The only exception occurs for linearly priced buyers; however, in this situation, they are much more sensitive to market running costs and decreasing the market welfare for lower market running cost values.

In sum, we can conclude from the simulation results that the logarithmic pricing function is better. From an economic perspective, this is also reasonable. Our market design was based on the Efficient Markets Hypothesis (EMH), where there is sufficient competition [45]. For the provider, the cost of S(t) data records < t* S(1) data records in a single transaction, so a monotonically increasing yet marginally decreasing pricing model is acceptable. For the buyer, purchasing the same amount of data, the tendency is to deal with the provider who offers a lower price. Faced with the same number of records S(t) and the same price for S(1), logarithmic pricing is much less expensive than linear pricing, and linearly priced providers will be eliminated. Thus, logarithmic pricing can better satisfy both sides of the price curve.

We can conclude that the logarithmic pricing function allows us to best reach the set objectives of maximizing data trading and market welfare.

Discussion

Motivated by the challenges that arise through a lack of decentralized data trading and data sharing among firms, and the potentials that exchanging and trading on-chain data would hold for car-related—and other—businesses, we asked RQ1*: How can we design a blockchain-based market for data trading?* Drawing on previous work and insights from the CarData project, we derived design requirements and market objectives, proposing a market model for data trading. PeBDaMa addresses the specific business needs (e.g., the need for a combinatorial data set from multiple providers, the need for adequate pricing) by utilizing the idiosyncrasies of a blockchains-based infrastructure (e.g., decentralized data collection and management, unique ownership of data goods). Building on market design research [1, 57], we proposed a market model and answered RQ1. To evaluate our market model and answer RQ2—*How should blockchain data be priced?*—we ran computational simulations, which allowed us to test how data trading and welfare can be maximized through the pricing function and to explore how the market behaves given the novel properties. Our results indicate that in a blockchain-based data market, data trading and welfare from it can be maximized by pricing data logarithmically. Yet this poses the risk of a sudden collapse if transaction fees become too high. These results answer our overarching research question and deliver several interesting insights for academics working on the challenges in the used-car market, blockchain and market design researchers, and practitioners designing and aiming to operate blockchain-based data markets, which we next discuss.

Addressing Current Problems of Data Buyers and Sellers

Given the potentials offered by exchanging and trading data, in recent years, platforms operated by an intermediary firm that enable many-to-many data trades have emerged. Yet, to date, they have not achieved widespread use; many have even failed [41, 76]. As noted, a prominent example is the Microsoft Azure Data Catalog. While some researchers argue that the lack of buyers is a major reason why such data markets have failed [57], following reference [41] we argue that the reasons are that it is hard to ascertain data quality on the buyer side and that it is hard to price data and control their use on the seller side. Building on a blockchain-based infrastructure enables data sellers to address these challenges. First, through the cryptographic mechanisms on the blockchain, changing the nature of data goods by creating unique digital ownership [56] allows one to find adequate pricing mechanisms for their valuable resources. Second, through the on-chain traceability, incorporating data owners' role allows them to return control of their data resources to those that generate and own the data [20]. Finally, a consortium of market managers/regulators with verifiable vetting of market participants is the model that underlies the consortium blockchain, with the ability to prove data provenance establishing the necessary trust in data quality [51, 97, 103]. Thus, for market designers and operators, we conclude that, rather than aiming to optimize market mechanisms for buyers [57], *establishing a verified infrastructure addresses the crux of the current challenges to data buyers and sellers, enabling one to create the truly decentralized data markets called for by the market*. Further, building on a blockchain-based decentralized infrastructure, following reference [29], we argue that the market objective should opt for true decentralization that favors neither party but, instead, incentivizes the equal participation of buyers and sellers and therefore optimizes the entire ecosystem.

Data on Blockchain and the Consequences for Data Pricing

The practicality of implementing a logarithmic pricing function in data markets is not straightforward. In typical databases, pricing functions that decrease with volume may incentivize reselling [86, 87]. Consider a scenario with logarithmic pricing: Buyers could acquire large volumes of data to resell to those seeking fewer records. Such reselling benefits buyers, but could diminish market and provider revenues.

Yet blockchain-based data markets offer a crucial, unique distinction. Here, records are verifiable and immutable, enhancing data's value and curbing arbitrage [55, 56]. Off-chain resale will not have the same value, especially when data authenticity and provenance are vital, such as in compliance-related processes or the used-car market [95]. We therefore suggest that market designers strongly consider logarithmic pricing for blockchain-verified data.

Our simulations revealed that a logarithmic pricing model yields the fewest transaction rejections across various demand functions: constant, linear, or logarithmic. Further, it maximizes market welfare. In conclusion, our primary recommendation for market designers of a blockchain-based data market is to *carefully consider the option of logarithmic pricing, even though it may seem counterintuitive.*

Changing Market Dynamics' Impacts on Market Elasticity and Transaction Fees

Finally, a caveat we observed from our simulation for market operators: As a result of the changing properties of data (i.e., unique control over and ownership of data) and their new possibilities for pricing (i.e., logarithmic pricing function), the market elasticity changes. Our simulations showed that the ability to price data logarithmically happens to perform fairly well for what is generally a typical demand for data requests in data markets [57, 72] and was also mentioned as a typical demand function in the car market by an interviewee: It poses the risk of a sudden market collapse if the transaction costs become too high. In today's data markets, a marginal increase in transaction fees leads to a marginal decrease in the number of transactions. In stark contrast, we observed that in a blockchain-based market, the number of transactions may suddenly cease owing to even small changes in the transaction pricing. This sensitivity may stem from the enhanced transparency associated with on-chain transactions.

This is an interesting and important finding, as transaction costs can become a danger in blockchain data markets. As observed for the Bitcoin blockchain and for other distributed technologies, a distributed and decentralized verification model is not inexpensive, and resulting high transaction costs may even endanger business models [25]. Thus, our primary recommendation for market designers and operators of blockchain data markets is *to be aware of the changing market elasticity and to carefully consider transaction fees so as to avoid a sudden market collapse*. For instance, we suggest that practitioners implement an index of the daily price elasticity of demand in the market. Once the index becomes higher, the transaction cost charges can be reduced appropriately so as to stabilize demand. Further, the transaction fee could be moderately increased so as to balance running costs and collect money for the construction of on-chain infrastructure when market demand is particularly robust or dropping.

Table 4. Comparison Between Established Data Trading Platforms and Blockchain-Based Data Trading Market Designs

Established data trading platforms	Blockchain-based data market design (logarithmic pricing)
Data quality is difficult to trust.	The consortium's permission mechanism safeguards data quality.
Controlling data use is hard for sellers.	Blockchain provides unique ownership of the data, and the use of the data is controlled by the seller.
Low pricing power for sellers and market bias toward buyers.	Decentralized market structure that incentivizes equal participation from both sides without favoring any side.
Data pricing must consider many issues, such as resale arbitrage, and may not always meet the requirement for maximum total market welfare.	Pricing models can be simplified, allowing for logarithmic pricing, maximizing total welfare, and somewhat reducing the number of declined transactions.
The number of transactions decreases at the margin as transaction costs increase.	Higher price elasticity of demand exists when special transaction prices make transactions suddenly disappear (market collapse).
	Market demand elasticity can be regulated by controlling the transaction fee.

Table 4 provides a comparison between established data trading platforms and blockchain-based data trading market designs.

Limitations

Our work has limitations that open avenues for future research. Given the state of our case project and the general lack of knowledge about blockchain data markets, the design of our market model was driven by market penetration and adoption objectives. However, as a project partner pointed out, it would be worth considering the special characteristics of the different phases of a blockchain data market's establishment in greater detail. While early on the objective most likely is to maximize the number of transactions, to incentivize participation in and adoption of the market, market designers and operators should be aware that each phase has different challenges and eventually will require the adaptation of the market mechanisms. Thus, we recommend that researchers study the different phases of the blockchain data market and their idiosyncrasies, and that they then design appropriate market mechanisms for each phase. Also, to keep the model minimalistic, we did not consider learning from the perspectives of market operators or buyers. It will be interesting to understand how the former can, by observing the market functioning, avoid an eventual collapse. Thus, researchers could introduce different behavioral strategies for market operators in the simulation, so as to explore the optimal strategies that they can adopt in order to avoid market collapse. Further, given the business-driven goals of firms that seek to understand whether or not their investments in the design and establishment of blockchain-based data markets will be worthwhile, we have focused on the pricing of data in such markets. While insecurities about financial returns are often a major hindrance in the development of such projects, many other key issues, such as security and privacy regulations of the decentralized infrastructure, as well as the handling of diverse quality of data in such markets, must first be resolved before such blockchain-based data markets can come alive.

Finally, this study is focused on market design research for a specific use case, thus setting many stringent conditions for pricing. In these conditions, we assumed that all data records have the same structure and value, which led our pricing model to consider pricing based on the number of data records, rather than assessing the value of single data records, which

distinguishes our work from most pricing research. However, in other scenarios, as the data provided by different data suppliers can vary in quality and richness, the assessment of a single data record's value becomes a key part of the pricing model. The studies we discuss in our literature review provide important references for further refining our research in the future. For instance, we could introduce an improved Shapley value method to assess single data records' value and can then apply this article's methodology to determine the final price of bundled data sets, and so on.

Conclusions

Data market design research has a major problem when the research should be implemented for a setting based on a real application. Our results generated by the design of the market model and its simulations provide useful foundations for academics and practitioners working on blockchain-based data markets for the car ecosystem [6, 7, 63, 98]. It also allows us to derive valuable insights and directions for the market design for both academics and practitioners who work on blockchain data markets and business models and who face similar challenges in other domains. The insights we present in this article can support their design efforts and provide key insights into early design decisions.

Despite these limitations and open research areas, our research provides a foundation for operationalizing a blockchain-based data market. It can be used to add extensions and run simulations to study agent behaviors and the market dynamics in different compensation/incentive schemes. Further, it can be used by practitioners as the basis to run usage case simulations and to determine optimal pricing strategies for certain application areas of verified car (or other) data. In our view, such blockchain-based decentralized data markets have the potential to totally change current assumptions about the use and trading of data goods. We trust that this study motivates other researchers to further explore and evaluate the potentials and value of blockchain-based data goods.

Acknowledgments

The authors thank the cardossier association and especially its member firms AMAG Import AG, Audatex Schweiz GmbH, Auto-I-Dat AG, and AXA for their time and participation in workshops and interviews, as well as their valuable inputs. We also thank Mirko Richter, who conducted his master's thesis during the course of this project and supported us in our data collection and analysis.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Ingrid Bauer-HÄnsel (ingrid.bauer-haensel@unisg.ch; corresponding author) is an assistant professor at the University of St. Gallen, Switzerland.. She holds an M.Sc. in information systems from the Copenhagen Business School, and a Ph.D. from the University of Zürich. Her research focuses on interorganizational systems, decentralized data platforms and markets, and token systems. She led the industry collaboration on this project.

Qianyu Liu (qianyu@ifi.uzh.ch) is a Ph.D. student in the Blockchain & Distributed Ledger Technologies group at the University of Zurich, Switzerland. Her research focuses on consortium blockchain and blockchain applications.

Claudio J.Tessone (claudio.tessone@uzh.ch) is a professor in Blockchain & Distributed Ledger Technologies at the University of Zurich and chairman of the UZH Blockchain Center. His main interests are complex socioeconomic and sociotechnical systems from an interdisciplinary perspective, aiming at unveiling links between microscopic agent behaviors, the rules they follow, and a system's emerging global properties. Blockchain-based systems, cryptocurrencies, and other digital tokens are pillars of his current activities. This includes crypto-economics, big data blockchain analytics and forensics, and the analysis of economic incentives, which are fundamental to multiple blockchain platforms. He also directs the interdisciplinary UZH Summer School: Deep Dive into Blockchain and co-directs the Certificate of Advanced Studies on Blockchain at the University of Zurich.

Gerhard Schwabe (schwabe@ifi.uzh.ch) is a professor in the Department of Informatics at the University of Zurich, where he leads the information management research group. He received his doctoral and postdoctoral education at the University of Hohenheim, Germany. Dr. Schwabe researches the intersection of collaborative technologies and information management. He has studied collaboration in commercial and government organizations at the granularity of dyads, small teams, large teams, organizations, communities, and social networks, frequently in collaboration with companies and public organizations. He has published numerous papers in major journals and conference proceedings, focusing on information systems and computer science.

ORCID

Ingrid Bauer-Hänsel http://orcid.org/0000-0001-8837-790X
Claudio J. Tessone http://orcid.org/0000-0001-7733-6221
Gerhard Schwabe http://orcid.org/0000-0002-0453-9762

References

1. Agarwal, A.; Dahleh, M.; and Sarkar, T. A marketplace for data: An algorithmic solution. In *Association for Computing Machinery, ACM Conference on Economics and Computation (EC '19)*, May 2019.
2. Akerlof, G. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84* (1970), 488–500.
3. Azcoitia, S.A.; and Laoutaris, N. *A Survey of Data Marketplaces and Their Business Models* (2022).
4. Bauer, I.; Parra-Moyano, J.; Schmedders, K.; and Schwabe, G. Multi-party certification on blockchain and its impact in the market for lemons. *Journal of Management Information Systems*, *39*, 2 (2022), 395–425.
5. Bauer, I.; Zavolokina, L.; Leisibach, F.; and Schwabe, G. Exploring blockchain value creation: The case of the car ecosystem. In *Proceedings of the 52nd Annual Hawaii International Conference on System Sciences*. Maui, HI, 2019, p. 10.
6. Bauer, I.; Zavolokina, L.; Leisibach, F.; and Schwabe, G. Value creation from a decentralized car ledger. *Frontiers in Blockchain*, *2* (2020).
7. Bauer, I.; Zavolokina, L.; and Schwabe, G. Is there a market for trusted car data? *Electronic Markets*, *30*, 2 (September 2019), 211–225.
8. Baumol, W.J.; and Bradiford, D.F. Optimal departures from marginal cost pricing. *American Economic Review*, *60*, (1970), 265–283.

9. Beck, R.; Avital, M.; Rossi, M.; and Thatcher, J.B. Blockchain technology in business and information systems research. *Business & Information Systems Engineering*, *59*, 6 (2017), 381–384.
10. Beck, R.; Czepluch Stemi, J.; Lollike, N.; and Malone, S. Blockchain—The gateway to trust-free cryptographic transactions. In *Twenty-Fourth European Conference on Information Systems*. Istanbul, Turkey, 2016, pp. 1–14.
11. Beck, R.; Müller-Bloch, C.; and King, J. Governance in the blockchain economy: A framework and research agenda. *Journal of the Association for Information Systems*, *19*, 10 (2018), 1020–1034.
12. Blundell, R.; Gu, R.; Leth-Petersen, S.; Low, H.; and Meghir, C. Durables and lemons: Private information and the market for cars. *US National Bureau of Economic Research,Working Paper Series*, 26281 (2019).
13. Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, *99*, Supplement 3 (2002), 7280–7287.
14. Bond, E.W. A Direct test of the "lemons" model: The market for used pickup trucks. *American Economic Review*, *72*, 4 (1982), 836–840.
15. Brown, D.J.; and Heal, G.M. Marginal vs. average cost pricing in the presence of a public monopoly. *American Economic Review*, *73*, 2 (1983), 189–193.
16. Buterin, V. *Ethereum white paper*, 2013. https://github.com/ethereum/wiki/wiki/White-Paper .
17. Buterin, V. DAOs, DACs, DAs and more: An incomplete terminology guide. *Ethereum Blog*, 2014. https://blog.ethereum.org/2014/05/06/daos-dacs-das-and-more-an-incomplete-terminology-guide.
18. Caruso. Caruso-Dataplace. 2022. https://www.caruso-dataplace.com.
19. Chen, Y. Blockchain tokens and the potential democratization of entrepreneurship and innovation. *Business Horizons*, *61*, 4 (2018), 567–575.
20. Cunningham, J.; and Ainsworth, J. Enabling patient control of personal electronic health records through distributed ledger technology. *Studies in Health Technology and Informatics*, *245* (2017), 45–48.
21. Dai, W.; Dai, C.; Choo, K.-K.R.; Cui, C.; Zou, D.; and Jin, H. SDTE: A secure blockchain-based data trading ecosystem. *IEEE Transactions on Information Forensics and Security*, *15* (2020), 725–737.
22. De Filippi, P.; and Loveluck, B. The invisible politics of Bitcoin: Governance crisis of a decentralized infrastructure. *Internet Policy Review*, *5*, 4 (2016).
23. Dozier, P.; and Saunders, C. The inter-organizational perspective in blockchain adoption within an ecosystem. In *Proceedings of the 28th European Conference on Information Systems*, An Online AIS Conference, 2020.
24. Drasch, B.J.; Fridgen, G.; Manner-Romberg, T.; Nolting, F.M.; and Radszuwill, S. The token's secret: the two-faced financial incentive of the token economy. *Electronic Markets* (March 2020).
25. Easley, D.; O'Hara, M.; and Basu, S. From mining to markets: The evolution of bitcoin transaction fees. *Journal of Financial Economics*, *134*, 1 (2019), 91–109.
26. Fan, Z.; Fang, H.; Zhou, Z.; et al. Improving fairness for data valuation in horizontal federated learning. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*, 2022, pp. 2440–2453.
27. Findlay, R. On W. Arthur Lewis' contributions to economics. *Scandinavian Journal of Economics*, *82*, 1 (1980), 62.
28. Gao, W.; Hatcher, W.G.; and Yu, W. A survey of blockchain: Techniques, applications, and challenges. In *27th International Conference on Computer Communication and Networks (ICCCN)* (2018), pp. 1–11.
29. Glaser, F. Pervasive decentralisation of digital infrastructures: A framework for blockchain enabled system and use case analysis. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.
30. Gozman, D.; Liebenau, J.; and Aste, T. A case study of using blockchain technology in regulatory technology. *MIS Quarterly Executive*, *19*, 1 (2020), 19–37.

31. Hansen, S. and Baroody, A.J. Electronic health records and the logics of care: Complementarity and conflict in the U.S. healthcare system. *Information Systems Research*, *31*, 1 (2020), 57–75.
32. Hevner, A.R. A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*, 2 (2007), 7.
33. Hevner, A.R.; March, S.T.; Park, J.; and Ram, S. Design science in information systems research. *MIS Quarterly*, *No. 1*, 28 (2004), 75–105.
34. Hicks, J. R. The foundations of welfare economics. *Economic Journal*, *49*, 196 (1939), 696–712.
35. Hu, D.; Li, Y.; Pan, L.; Li, M.; and Zheng, S. A blockchain-based trading system for big data. *Computer Networks*, *191*, (May 2021), 107994.
36. Jackson, J.E.; and Xu, X. Does scarcity add value in influencing consumers in the try-before-you-buy model? *International Journal of Electronic Commerce*, *26*, 1 (January 2022), 25–48.
37. Jaiman, V.; and Urovi, V. A consent model for blockchain-based health data sharing platforms. *IEEE Access*, *8* (2020), 143734–143745.
38. Jensen, T.; Hedman, J.; and Henningsson, S. How TradeLens delivers business value with blockchain technology. *MIS Quarterly Executive*, *18*, 4 (2019), 221–243.
39. Kaiser, C.; Stocker, A.; Viscusi, G.; Fellmann, M.; and Richter, A. Conceptualising value creation in data-driven services: The case of vehicle data. *International Journal of Information Management*, *59* (2021), 102335.
40. Koutris, P.; Upadhyaya, P.; Balazinska, M.; Howe, B.; and Suciu, D. Toward practical query pricing with QueryMarket. In *Proceedings of the 2013 International Conference on Management of Data—SIGMOD '13*. New York: ACM Press, 2013, p. 613.
41. Koutroumpis, P.; Leiponen, A.; and Thomas, L.D.W. Markets for data. *Industrial and Corporate Change*, *29*, 3 (2020), 645–660.
42. Lee, J.S.; Pries-Heje, J.; and Baskerville, R. Theorizing in design science research. In H. Jain, A. P. Sinha and P. Vitharana (eds.), *token*. Berlin: Springer, 2011, pp. 1–16.
43. Liang, T.-P.; Kohli, R.; Huang, H.-C.; and Li, Z.-L. What drives the adoption of the blockchain technology? A fit-viability perspective. *Journal of Management Information Systems*, *38*, 2 (2021).
44. Lindman, J.; Rossi, M.; and Virpi, K.T. Opportunities and risks of blockchain technologies in payments—A research agenda. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.
45. Lo, A.W. *Efficient Markets Hypothesis* (2007).
46. López, D.; and Farooq, B. A multi-layered blockchain framework for smart mobility data-markets. *Transportation Research Part C: Emerging Technologies*, *111*, (2020), 588–615.
47. Ma, C.; Li, J.; Ding, M.; et al. When federated learning meets blockchain: A new distributed learning paradigm. 2021. http://arxiv.org/abs/2009.09338.
48. Macal, C.M.; and North, M.J. Agent-based modeling and simulation. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*. IEEE, Austin, TX, 2009, pp. 86–98.
49. March, S.T.; and Smith, G.F. Design and natural science research on information technology. *Decision Support Systems*, *15*, 4 (1995), 251–266.
50. Marchand, M.G. The economic principles of telephone rates under a budgetary constraint. *Review of Economic Studies*, *40*, 4 (1973), 507–515.
51. Marella, V.; Upreti, B.; Merikivi, J.; and Tuunainen, V.K. Understanding the creation of trust in cryptocurrencies: The case of Bitcoin. *Electronic Markets*, *30*, 2 (2020), 259–271.
52. Mattioli, M. Disclosing big data. *Minnesota Law Review*, *99* (2014), 534–584.
53. McMahan, H.B.; Moore, E.; Ramage, D.; Hampson, S.; and Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. 2017. http://arxiv.org/abs/1602.05629 .
54. McMahan, H.B.; and Ramage, D. Federated learning: Collaborative machine learning without centralized training data. 2017. https://ai.googleblog.com/2017/04/federated-learning-collaborative.html .
55. Miscione, G.; Goerke, T.; Klein, S.; Schwabe, G.; and Ziolkowski, R. From authentication to "Hanseatic governance": Blockchain as organizational technology. 2019.

56. Miscione, G.; Richter, C.; and Ziolkowski, R. Authenticating deeds/organizing society: Considerations for blockchain-based land registries. In W. De Vries (ed.), *Responsible and Smart Land Management Interventions: An African Context*. Boca Raton, FL: CRC Press, Taylor & Francis, 2020, pp. 133–147.
57. Moor, D.; Seuken, S.; Grubenmann, T.; and Bernstein, A. The design of a combinatorial data market. Technical report, University of Zurich (2019), 41.
58. Myers, M.D.; and Newman, M. The qualitative interview in IS research: Examining the craft. *Information and Organization*, *17*, 1 (2007), 2–26.
59. Naerland, K.; Müller-Bloch, C.; Beck, R.; and Palmund, S. Blockchain to rule the waves—Nascent design principles for reducing risk and uncertainty in decentralized environments. Seoul, South Korea, 2017.
60. Nakamoto, S. Bitcoin P2P e-cash paper. *The Cryptography Mailing List*, 2008. http://www.metzdowd.com/pipermail/cryptography/2008-October/014810.html .
61. Ng, Y.-K.; and Weisser, M. Optimal pricing with a budget constraint—The case of the two-part tariff. *Review of Economic Studies*, *41*, 3 (1974), 337.
62. Nguyen, D.C.; Ding, M.; Pham, Q.-V.; et al. Federated learning meets blockchain in edge computing: Opportunities and challenges. 2021. http://arxiv.org/abs/2104.01776.
63. Notheisen, B.; Cholewa, J.B.; and Shanmugam, A.P. Trading real-world assets on blockchain: An application of trust-free transaction systems in the market for lemons. *Business & Information Systems Engineering*, *59*, 6 (2017), 425–440.
64. Ocean Protocol Foundation. Ocean Protocol: A decentralized substrate for AI data & services. Technical whitepaper. Ocean Protocol Foundation1 with BigchainDB GmbH2 and Newton Circus (DEX Pte. Ltd.)3, Version 2019-MAR-05 (2019).
65. Ostern, N.K. Blockchain in the IS research discipline: A discussion of terminology and concepts. *Electronic Markets* (December 2019).
66. Parra-Moyano, J.; Schmedders, K.; and Pentland, A. What managers need to know about data exchanges. *MIT Sloan Management Review*, *61*, 4 (2020), 39–44.
67. Peck, M.E. Blockchain world—Do you need a blockchain? This chart will tell you if the technology can solve your problem. *IEEE Spectrum*, *54*, 10 (2017), 38–60.
68. Peffers, K.; Tuunanen, T.; Rothenberger, M.A.; and Chatterjee, S. A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*, 3 (2007), 45–77.
69. Pei, J. A Survey on data pricing: From economics to data science. *IEEE Transactions on Knowledge and Data Engineering*, *34*, 10 (2022), 4586–4608.
70. Risius, M.; and Spohrer, K. A blockchain research framework: What we (don't) know, where we go from here, and how we will get there. *Business & Information Systems Engineering*, *59*, 6 (2017), 385–409.
71. Rossi, M.; Mueller-Bloch, C.; Thatcher, C.; Bennett, J.; and Beck, R. Blockchain research in information systems: Current trends and an inclusive future research agenda. *Journal of the Association for Information Systems*, *20*, 9 (2019), 1388–1403.
72. Saichev, A.; Malevergne, Y.; and Sornette, D. *Theory of Zipf's Law and Beyond*. Berlin: Springer, 2010.
73. Saldaña, J. *The Coding Manual for Qualitative Researchers*. Los Angeles, CA: Sage, 2009.
74. Sarker, S.; Henningsson, S.; Jensen, T.; and Hedman, J. Blockchain as a strategy for combating corruption in global shipping: An interpretive case study. *Journal of Management Information Systems*, *38*, 2 (2021), 338–373.
75. Schallbruch, M.; Schweitzer, H.; and Wambach, A. Europa stutzt die Digitalkonzerne. *Frankfurter Allgemeine Zeitung (FAZ)*, 2021. https://www.faz.net/aktuell/wirtschaft/europa-stutzt-die-digitalkonzerne-kampf-gegen-monopolstellungen-17158280.html.
76. Schomm, F.; Stahl, F.; and Vossen, G. Marketplaces for data: An initial survey. *ACM SIGMOD Record*, *42*, 1 (2013), 15–26.
77. Schwill, F.C. Towards decentralized and privacy-preserving data marketplaces to unlock data for AI: An examination of Ocean Protocol. White paper (2021).

78. Seebacher, S.; and Schüritz, R. Blockchain—Just another IT implementation? A comparison of blockchain and interorganizational information systems. In *Proceedings of the 27th European Conference on Information Systems*. Stockholm–Uppsala, Sweden, 2019.
79. Shapiro, C.; and Varian, H. *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press, 1999.
80. Spychiger, F.; Zavolokina, L.; and Schwabe, G. Incentivizing data quality in blockchain-based systems—The case of the digital cardossier. *ACM Distributed Ledger Technologies* (2022).
81. Subramanian, H. Decentralized blockchain-based electronic marketplaces. *Commun. ACM*, *61*, 1 (2018), 78–84.
82. Sunyaev, A.; Kannengießer, N.; Beck, R.; et al. Token economy. *Business & Information Systems Engineering* (February 2021).
83. Tadeneke, A. case study on data markets in India and Japan show what is possible. Public Engagement, World Economic Forum, ated@weforum.org, 2021. https://www.weforum.org/press/2021/08/case-study-on-data-markets-in-india-and-japan-show-what-is-possible.
84. Tasca, P.; and Tessone, C.J. A taxonomy of blockchain technologies: Principles of identification and classification. *Ledger*, *4* (February 2019).
85. The Economist. The juicy market for lemons. Can you buy a good second-hand car? *The Economist* (September 2019).
86. Toral, R.; Tessone, C.J.; and Lopes, J.V. Collective effects induced by diversity in extended systems. *European Physical Journal Special Topics*, *143*, 1 (2007), 59–67.
87. Varian, H.R. *Pricing Information Goods*. Ann Arbor: University of Michigan, 1995.
88. Venable, J.; Pries-Heje, J.; and Baskerville, R. A comprehensive framework for evaluation in design science research. In K. Peffers, M. Rothenberger, and B. Kuechler (eds.), *Design Science Research in Information Systems. Advances in Theory and Practice*. Berlin: Springer, 2012, pp. 423–438.
89. Viswanathan, S.; and Anandalingam, G. Pricing strategies for information goods. *Sadhana*, *30*, 2 (2005), 257–274.
90. Wan, P.K.; Huang, L.; and Holtskog, H. Blockchain-enabled information sharing within a supply chain: A systematic literature review. *IEEE Access*, *8* (2020), 49645–49656.
91. Wang, T.; Rausch, J.; Zhang, C.; Jia, R.; and Song, D. A principled approach to data valuation for federated learning. In Q. Yang, L. Fan and H. Yu (eds.), *Federated Learning: Privacy and Incentive*. Cham: Springer International, 2020, pp. 153–167.
92. WEF. Data for common purpose: Enabling Colombia's Transition to a data-driven economy. In *Collaboration with PwC Colombia and the Centre for the Fourth Industrial Revolution Colombia*, 2021. https://www.weforum.org/whitepapers/data-for-common-purpose-enabling-colombia-s-transition-to-a-data-driven-economy.
93. Wilensky, U.; and Rand, W. *An Introduction to Agent-Based Modeling*. Cambridge, MA: MIT Press, 2015.
94. Wingreen, S.C.; Kavanagh, D.; Dylan-Ennis, P.; and Miscione, G. Sources of cryptocurrency value systems: The case of Bitcoin. *International Journal of Electronic Commerce*, *24*, 4 (2020), 474–496.
95. Wörner, D.; Bomhard, T.V.; Schreier, Y.-P.; and Bilgeri, D. The Bitcoin ecosystem: Disruption beyond financial services? *European Conference on Information Systems (ECIS)*, (2016).
96. Zavolokina, L.; Miscione, G.; and Schwabe, G. Buyers of "lemons": How can a blockchain platform address buyers' needs in the market for "lemons"? *Electronic Markets*, *30* (2020), 227–239.
97. Zavolokina, L.; Zani, N.; and Schwabe, G. Designing for trust in blockchain platforms. *IEEE Transactions on Engineering Management* (2020), 1–15.
98. Zavolokina, L.; Ziolkowski, R.; Bauer, I.; and Schwabe, G. Management, governance, and value creation in a blockchain consortium. *MIS Quarterly Executive*, *19*, 1 2020), 1–17.
99. Zhang, W.; Wei, C.P.; Jiang, Q.; Peng, C.H.; and Zhao, J.L. Beyond the block: A novel blockchain-based technical model for long-term care insurance. *Journal of Management Information Systems*, *38*, 2 (2021), 374–400.

100. Zhang, X.; Zha, X.; Zhang, H.; and Dan, B. Information sharing in a cross-border e-commerce supply chain under tax uncertainty. *International Journal of Electronic Commerce*, *26*, 1 (2022), 123–146.
101. Zheng, Z.; Xie, S.; Dai, H.; Chen, X.; and Wang, H. An overview of blockchain technology: Architecture, consensus, and future trends. In *Big Data (BigData Congress), 2017 IEEE International Congress on*. IEEE, 2017, pp. 557–564.
102. Zwass, V. Editor's introduction. *International Journal of Electronic Commerce*, *22*, 4 (2018), 477–478.
103. Zwass, V. Editorial introduction. *Journal of Management Information Systems*, *38*, 2 (2021), 277–281.
104. Myerson, Roger B. "Optimal Auction Design." Mathematics of Operations Research, vol. 6, no. 1, 1981, pp. 58–73. JSTOR, http://www.jstor.org/stable/3689266. Accessed 4 Jan. 2024.