Discovering Business Models of Data Marketplaces

Michael Fruhwirth
Know-Center GmbH
mfruhwirth@know-center.at

Michael Rachinger
Graz University of Technology
michael.rachinger@tugraz.at

Emina Prlja
Graz University of Technology
emina.prlja@student.tugraz.at

Abstract

The modern economy relies heavily on data as a resource for advancement and growth. Data marketplaces have gained an increasing amount of attention, since they provide possibilities to exchange, trade and access data across organizations. Due to the rapid development of the field, the research on business models of data marketplaces is fragmented. We aimed to address this issue in this article by identifying the dimensions and characteristics of data marketplaces from a business model perspective. Following a rigorous process for taxonomy building, we propose a business model taxonomy for data marketplaces. Using evidence collected from a final sample of twenty data marketplaces, we analyze the frequency of specific characteristics of data marketplaces. In addition, we identify four data marketplace business model archetypes. The findings reveal the impact of the structure of data marketplaces as well as the relevance of anonymity and encryption for identified data marketplace archetypes.

1. Introduction

The amount of available data has increased exponentially in recent years and, when combined with new possibilities in the field of data analysis, machine learning and storage technologies, data can be leveraged more frequently for economic benefit. Thus, data and analytics have the potential to transform a firm's business models [19, 25, 54]. While the monetization of data is an essential focus of organizations and even industries today, few market mechanisms exist to exchange and price such data sets and match buyers to sellers [1]. Data marketplaces seem to be one approach that can be taken to overcome those challenges [1, 33]. In order to be sustainable, such data marketplaces need an appropriate business model.

Little research has been conducted in the emerging field of data marketplaces in general [48] and business models of data marketplaces in particular [38]. Due to the special characteristics of data as economic goods, as compared to tangible goods, such as a lack of established rules and market mechanisms for pricing data goods, data marketplaces have distinct characteristics that differentiate them from other electronic marketplaces. In this research, we aimed to identify characteristics of data marketplaces by addressing the following research question: *"What are the characteristic elements of data marketplaces from a business model perspective?"*

In order to identify those characteristics, a taxonomy was developed. Taxonomies are used to classify objects of interest in the domain of interest and help understand the complexity of the domain and its existing or possible concepts [30]. To develop the taxonomy, we used the clearly structured and well-tested method described in [30]. We conducted a literature review to identify conceptual characteristics of data marketplaces and used a final set of 20 cases of data marketplaces to empirically revise our taxonomy. We identified 16 key dimensions that could be used to distinguish and explain the dimensions and characteristics of data marketplaces from a business model perspective. The developed taxonomy contributes to the existing body of knowledge by establishing a common understanding of data marketplaces from a business model perspective.

The rest of the paper is organized as follows: In section two, we provide the theoretical background and related work on business models, data as central business resource, platform and electronic marketplace business models as well as data marketplaces, which form the basis for our research. Section three describes our research methodology, detailing the steps of taxonomy development, including how we collected and analyzed our empirical data. In the next section we present the individual elements of our taxonomy, a frequency analysis, four archetypes of data marketplaces as well as four cases illustrating the found archetypes. We close the paper by discussing the implications of our research, reflecting its limitations and describing possible directions for further research.

HĮCSS

2. Theoretical background

2.1. Business models

The concept of business models has gained momentum in recent years in several disciplines, such as Information Systems [7, 49], Strategic Management [26, 56] as well as Technology and Innovation Management [5, 10]. Companies need well-developed business models to *"commercialize new ideas and technologies"* [9] and to deliver and capture value [47]. Developing business model components and frameworks is a popular approach that can be taken to describe and analyze business models. More specifically, a *"business model framework describes the compositional elements what a business model is made of"* [43]. A component-based view is subsequently present in many ways of understanding the term business model [51]. The contents of a business model are often thereby characterized by subordinated elements [51]. More precisely, business models can be applied to describe how organizations create, deliver and capture value through the value proposition [34, 47]. These subordinate elements subsequently provide the meta- characteristic for our taxonomy that are suitable for use to analyze data marketplaces using the business model as a viewing lens.

2.2. Data as the main assets in business models

Data and information are strategic resources in today's digital businesses [6, 27]. Data enable firms to improve processes and decision-making, offer new services and even new business models [20, 53]. This leads to the paradigm that data assets and information-based offerings are exchanged for legal tender [52]. Thus, through the emergence of a data ecosystem, data do not only represent an asset that allows companies to improve products or services, but have become products in and of themselves [8, 39].

However, data have special characteristics as compared to tangible goods. Data are easy to transport, share, or copy, and they can be equally used in any location or environment, which gives them advantages over any other product [24]. Compared to tangible goods, there are no established rules and market mechanisms for pricing data assets [17, 28]. In addition, it is often difficult for data buyers to evaluate data assets before purchasing and fully disclosing and accessing them, a conundrum that is also known as *Arrow's Paradox* [3, 40]. Furthermore, organizations are often not willing to share valuable data with other stakeholders due to potential risks or a lack of trust [11]. As when trading other goods, data that are being offered on marketplaces should come with corresponding metadata, describing their amounts, origins, characteristics and other information that the buyer should know prior to conducting the purchase [39].

Data-driven business rely on data as a key resource [20]. Sources of data for data-driven business models can be *internal* (i.e., obtained from company's internal IT systems or self-generated sensor data) or *external* (i.e., data purchased from external data providers or obtained for free from publicly available data sources) [20]. Thus, companies purchase data goods to support their data-driven businesses (e.g., for training machine learning models [1]). Likewise, companies also have the opportunity to monetize their internal valuable data assets by selling them to other data-driven businesses [32]. Thus, data-driven businesses require market mechanisms and exchange infrastructure to acquire and monetize valuable data assets.

2.3. Electronic marketplaces and platforms

Digital platforms, which are often referred to as "electronic marketplaces" or "multilateral marketplaces", are businesses that enable and support interactions between distinct but interdependent groups of users (customers and suppliers). These groups perform exchanges of goods by using pricing strategies [18, 31]. The platform acts as facilitator of these interactions, and as participants co-create value between each other, a "network effect" is created. This means that a good or a service acquires more value to its user as more users adopt it [31, 37]. Network effects create self-reinforcing mechanisms that lead to market leadership, a large customer base economy of scale and the establishment of boundaries for other players [8, 16]. Participants in a marketplace do not necessarily represent two different groups of users, but can take both the roles of buyer and seller [46]. When data are subsequently traded as economic goods in electronic marketplaces, data marketplaces emerge as a type of marketplace with distinct characteristics.

2.4. Data marketplaces

Data marketplaces are electronic platforms that facilitate the exchange of data [41]. Due to the need for businesses to obtain appropriate information and the maturation of the data market ecosystem, the popularity of data marketplaces has grown in recent years [8]. As the number of data marketplace players joining the market increases, companies have more opportunities to use external data to improve their business and explore new revenue opportunities by reselling the data they collected internally [8, 29].

A data marketplace ecosystem consists of data providers, data buyers, third-party service providers and

a marketplace owner [38]. Data providers offer their own data on a marketplace; allow their data to be queried by data buyers and expect to obtain revenue by selling data. Data buyers are participants who are interested in buying the datasets they need and display a positive willingness to pay for data [23]. Data buyers use purchased data to support decision-making processes [12] or build new services and business models. Third-party service providers can provide applications or algorithms that add value to data assets. The data marketplace provider collects and hosts data from data providers and sells data to data buyers [38].

According to [42], an electronic marketplace is a data marketplace if data trading is the main value proposition of the business model. Marketplace participants have to be able to upload, browse, download, buy, and/or sell machine-readable data. Consequently, this excludes the services that only offer links to data locations, without hosting the data. Moreover, data have to be hosted by the providers who clarify the origin of the data. [8] emphasize additionally that data marketplaces have to be digital platforms and not only a data repository or a cloud service provider.

Encouraged by the definition of electronic marketplaces provided by [36], that involves characteristics of both *buying* and *selling*, open data marketplaces were not included in the scope of this research because they lacked a profit-oriented nature [57]. Open data portals or government agencies and non-government organizations that provide data free of charge were also excluded from the scope of data marketplaces, since they do not provide any market mechanism of buying and selling and the published data are only a side effect of their purpose in general.

Although the body of literature on data marketplaces has grown in recent years, little research has been conducted on the characteristics of data marketplaces from the perspective of business models.

3. Research methodology

In order to identify the characteristics of data marketplaces, the taxonomy-building approach proposed by [30] was chosen. Taxonomy building is an approach commonly used to classify, clarify, understand and systematically analyze complex problems or domains. The approach of [30] involves combining knowledge obtained from the literature and directly analyzing objects of interest. It is generally a well-accepted and frequently used method in the area of business models for information systems, such as car sharing business models [35], digital business models [6] or FinTech business models [15]. In the following section, we describe how we developed a taxonomy of dimensions and elucidated the characteristics of data

marketplaces from a business model perspective by following the seven-step process for taxonomy development proposed by [30] (see Figure 1): As part of this process, the meta-characteristic first has to be defined, which is a basis and limitation for discovery of dimensions and characteristics and which are derived accordingly. Furthermore, the ending conditions have to be established during this process, because it has an iterative nature, which requires constraints for termination. Third, the process can continue along one of two paths: a conceptual-to-empirical approach, building the taxonomy from relevant literature, or an empirical-to-conceptual approach, building the taxonomy from an investigation of analysis objects. After each approach, the ending conditions are checked: If the ending conditions are not met, an additional empirical-to-conceptual or conceptual-to-empirical approach. If the ending conditions are met, the process terminates, and the taxonomy development is considered as finished.
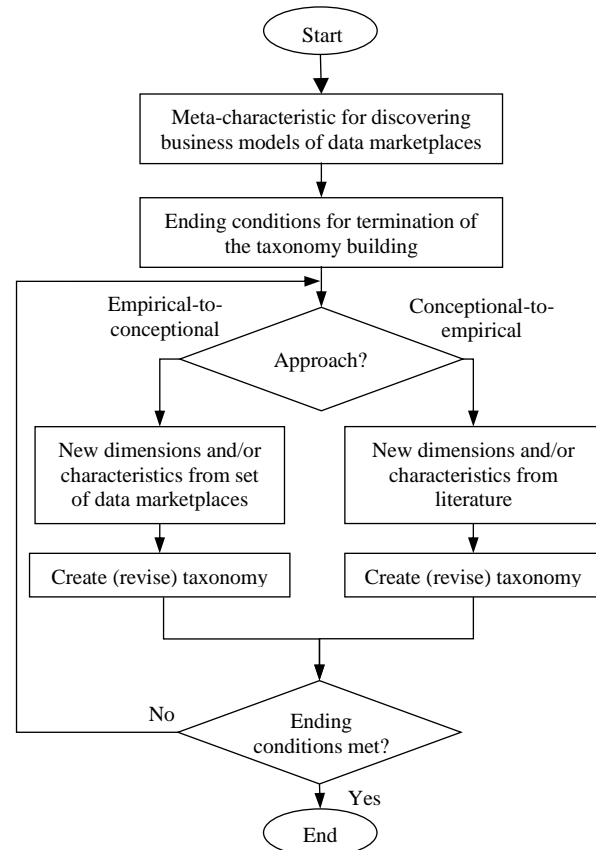
Figure 1. Applied research process in accordance with [30]

Define meta-characteristic: As the goal of this research was to identify characteristics of data marketplaces from a business model perspective, we

defined the basic elements of a business model (value proposition, value delivery, value creation and value capture [34, 47]) as the meta-characteristic.

Determine ending conditions: We followed the ending conditions suggested by [30] in terms of the objective and subjective ending conditions. The conditions are presented in Table 1, together with the taxonomy evolvement throughout the iterations.

Select approach: As part of our taxonomy development process, we conducted two conceptual-to-empirical iterations and five subsequent empirical-to-conceptual iterations. In the following section, we describe the activities in both approaches.

Conceptual to empirical (c2e): In the first two iterations, a conceptual-to-empirical approach was applied. We integrated relevant characteristics from the existing literature on taxonomies of platform business models and marketplaces [45, 46] as well as from previous work on data marketplaces [17, 22, 29, 40, 41].

Empirical to conceptual (e2c): In the next iterations, we classified the business models of data marketplaces from empirical data. To systematically identify the relevant objects for the inductive iterations, we adopted a rigorous procedure to identify relevant publications for a literature review, as suggested by [50]. We searched for data marketplaces using the Google search engine, using the browser in incognito mode to avoid carry-over effects from previous search. We applied the following keywords during our search: "data marketplace," "data market," "data trading platform," "data platform," "buying data" and "data brokers." Moreover, data marketplaces that had already been surveyed in the inspected literature, described in the background section, were included into our database. This search process led us to identify a total set of 58 data marketplaces. We drew on information from company websites, white papers and news articles that mentioned a data marketplace. Companies that provided insufficient information were omitted from the database. If possible, we also created an account for each data marketplace to observe its functionalities and offerings. Only objects which were available in the English language were considered for inclusion in the set of analysis objects. To create a representative set of platforms, the initial set of objects was filtered in two iterations: Prior to the taxonomy development, objects were excluded if they did not fulfill our definition of data marketplaces (see section 2.4). During the taxonomy development, an additional set of objects was excluded if no sufficient public information was available about the data marketplace (i.e., we did not contact marketplace operators by direct e-mail correspondence or phone calls to obtain further information), if technical failures and location issues (e.g., unsupported area) of the data marketplace

prevented information retrieval, or if the platform was under construction or in testing phase. (The list of identified data marketplaces and the criteria for exclusion is available on request from the authors.) After applying our inclusion and exclusion criteria, we obtained a final set of 20 data marketplaces.

Checking ending conditions was done after every iteration, until every condition had been satisfied once, which terminated the analysis. This process is shown in Table 1.

Table 1. Ending conditions in each iteration

	Ending condition	Iteration						
		1	2	3	4	5	6	7
		(c2e)		(e2c)				
Objective	All objects from representative sample were examined					x	x	x
	No object was merged or split in the last iteration	x	x	x	x	x	x	x
	Every characteristic of every dimension describes at least one object							x
	No new dimensions or characteristics were added in the last iteration						x	x
	No dimensions or characteristics were merged or split in the last iteration		x					x
	Every dimension is unique	x	x	x	x	x	x	x
	Every characteristic is unique within its dimension	x	x			x	x	x
	Each cell is unique and is not repeated	x	x	x	x	x	x	x
Subjective	Concise – taxonomy is meaningful but not overwhelming							x
	Robust – dimensions and characteristics differentiate sufficiently					x	x	x
	Comprehensive – all dimensions of interest identified and possible to reuse					x	x	x
	Extendible – easy to add new dimensions/characteristics	x	x	x	x	x	x	x
	Explanatory – dimensions / characteristics can describe an object from the domain		x	x	x	x	x	x

The last step taken during our research involved using the developed taxonomy and analysis results to identify possible patterns within analyzed data marketplaces. The identification of patterns was performed following the guidelines suggested by [55]. This involved recognizing similarities and dissimilarities within the cases and, eventually, a separation of mutually similar groups between them. As suggested in the guidelines, the number and variety of cases for this analysis increased the validity of the results. The analyzed data marketplaces were consequently grouped according to their characteristics, and meaningful archetypes of data marketplaces were identified. This process was guided by adopting the typology development guidelines provided by [14].

4. Results

4.1. Taxonomy of data marketplaces

Based on the methodology laid out in section 3, a taxonomy system to characterize data marketplaces was established using business models as a lens. The taxonomy consisted of dimensions that were evaluated based on identified characteristics of data marketplaces. Table 2 gives an overview of the identified categories as well as the respective characteristics.

The categories laid out in Table 2 were structured along the basic elements of a business model (value proposition, value delivery, value creation and value capture [34, 47]). A detailed description of the elements of the taxonomies and their origins can be found in Table 4 in the Appendix. Characteristics that did not occur in the final dataset of twenty data marketplaces were not included in the taxonomy. The excluded characteristics were "Government" as a data origin, "Web interface" as a data access type, "XML" and "RDF" as data output types, "complete access" as pre-purchase test-option, the domains of "scientific" and "social media," "C2C" and "C2B" marketplace types and the "Free" and "two-part tariff" pricing model. In addition, if no information on a dimension was available, the characteristic "No info" was introduced. The final dataset of twenty marketplaces was used to analyze the occurrence frequency of the identified characteristics. Table 2 also shows the results of the analysis. In the following section, the dimensions and characteristics are described along the meta-characteristic value creation, value proposition, value delivery and value capturing.

Value creation in data marketplaces. The *platform infrastructure* dimension refers to the architecture of a data marketplace as a multi-sided platform. In a centralized data marketplace, data is offered via a centralized location, whereas in a decentralized data marketplace data assets remain at the data provider, using e.g. a blockchain [22]. The results indicate that approximately two-thirds of the investigated platforms operated in a centralized manner while only one-third operated in a decentralized manner. The *data origin* dimension specifies where the offered data comes from (i.e., internet, self-generated, user-generated, community, government or authority) [40]. Self-generated data from private sources represented the most prominent source of data. The *review system* dimension describes if data assets are evaluated by users or the marketplace [45, 46]. Interestingly, eighteen of the twenty data marketplaces either did not have a review system or offered no information on a review system.

Value proposition in data marketplaces: The *privacy* dimension indicates if data marketplaces offer privacy-preserving mechanisms as part of their value proposition to increase the data providers' willingness to share their data while preserving privacy and confidentiality of the data. Half of the investigated data marketplaces offered a way to protect privacy guarantees through encryption or anonymization of data assets. Although half of the data marketplaces did not provide any information on privacy, this dimension was maintained due to its relevance. Fourteen of the twenty data marketplaces (i.e., the majority), guarantee the quality of data assets as part of the value proposition, which is indicated in the *data quality guarantee* dimension. The *time relevancy* dimension describes if static or dynamic (i.e., regularly updated) data sets are offered by the marketplace [40]. Seventeen of the investigated twenty data marketplaces also offered both regularly updated datasets in combination with static datasets or exclusively regularly updated datasets. The *pre-purchase testability* dimension refers to Arrow's Paradox (see section 2.2), if data assets can be accessed prior to purchase to evaluate the value of the data [40]. Only seven of the twenty data marketplaces offered restricted access to data assets prior to purchase, whereas the majority did not offer pre-purchase testability of their data sources.

Value delivery in data marketplaces: The *data output type* describes the technical format of the data asset (e.g., CSV/XLS, JSON or report) [41]. The *type of access* dimension describes how data assets can be accessed (i.e., via API, download or specialized software). Fifteen of the twenty data marketplaces relied on APIs or downloads to offer access to data. The *additional purchase support* dimension indicates if additional services (e.g., for data analysis) are offered for free of with an additional fee. More than half of the investigated data marketplaces offered additional purchase support, however, predominantly for an extra charge. The *domain* dimension specifies about what the data asset contains information (e.g., financial or sensor data) [17, 41]. No specific focus regarding the domain of offered datasets was identified in the analysis of data from the investigated twenty data marketplaces. The *marketplace participant* dimension refers to the stakeholders that are matched via a data marketplace [46]. While almost half of the data marketplaces also took a B2B focus or no focus at all, three of the twenty data marketplaces (i.e., only a minority) used a C2B model. The *smart contract with blockchain* dimension describes if smart contract as a privacy-preserving and safe payment method is implemented to enforce trust. Almost half of the investigated data marketplaces (nine out of twenty) offered smart contracts with a blockchain.

Table 2. Identified dimensions and characteristics of data marketplaces

	Dimension	Characteristics					
Value Creation	Platform infrastructure	Centralized (13/20)			Decentralized (7/20)		
	Data origin	Internet (1/20)	Self-generated (10/20)	User (3/20)	Community (2/20)	Author-ity (4/20)	
	Review System	User reviews (2/20)	Reviews by marketplace (2/10)		None (9/20)	No info (7/20)	
Value Proposition	Privacy	Anonymized (6/20)	Encrypted (2/20)		Both (2/20)	No info (10/20)	
	Data quality guarantee	Yes (14/20)			No (6/20)		
	Time relevancy	Static (3/20)	Dynamic (11/20)		Both (6/20)		
	Pre-purchase testability	None (12/20)	Restricted access (7/20)		No info (1/20)		
Value Delivery	Data output type	CSV/XLS (6/20)	JSON (4/20)	Report (1/20)	Multiple options (4/20)	No info (5/20)	
	Type of access	API (7/20)	Down-load (4/20)	Specialized Software (3/20)	API/Download (4/20)	No info (2/20)	
	Additional purchase support	With additional costs (8/20)	Included in price (3/20)		No (9/20)		
	Domain	All / Any (5/20)	Finance (2/20)	Geo (2/20)	Address (2/20)	Sensor (4/20)	Per-sonal (5/20)
	Marketplace participants	B2B (9/20)	C2B (3/20)		Any (8/20)		
	Smart contract with blockchain	Yes (9/20)			No (11/20)		
Value Capture	Pricing model	Usage based (7/20)	Package pricing (3/20)	Flat fee tariff (5/20)	Freemium (4/20)	No info (1/20)	
	Price discovery	Fixed prices (11/20)	Set by sellers (6/20)	Set by byers (1/20)	Auction (1/20)	Negoti-ation (1/20)	
	Payment currency	Cypto (6/20)	Fiat (13/20)		Both (1/20)		

Value Capture in data marketplaces: The *pricing model* dimension indicates the strategy of a data marketplace for gaining profit [29, 41]. The most prominent pricing models used were usage-based models (used by eight out of twenty data marketplaces), a flat fee tariff (used by five data marketplaces) and the Freemium-model (used by four out of twenty data marketplaces). Furthermore, the *price discovery* dimension describes how the price of a data set is determined prior to transaction [45, 46]. Eleven out of twenty data marketplaces relied on fixed prices, while six data marketplaces relied on prices set by sellers. The *payment currency* dimension indicates in which form a data marketplace handles payments. Fiat money was the most prominent payment currency, while only six offered payments via cryptocurrency. Only one data marketplace offered both payments with cryptocurrency and fiat money.

4.2. Archetypes of data marketplaces

As described in section 3, we explored patterns in the distinct characteristics of the investigated data marketplaces to identify archetypes of data marketplaces. We looked for meaningful similarities and dissimilarities within the cases by comparing the different dimensions of our taxonomy. The four archetypes differed from each other in the dimension of platform infrastructure (centralized vs. decentralized). Centralized data marketplaces also differed if they provide encryption and smart contracts or not. Decentralized data marketplaces differed if they were offering self-or user generated data, obtained from the personal domain, operated in a C2B context and if data could be accessed via API/download or by use of specialized software. The final sample of twenty data marketplaces was allocated to one of four archetypes. Table 3 provides an overview of the identified archetypes and lists the representative data marketplaces as well as their respective main characteristics.

Table 3. Illustrative examples of data marketplace business model archetypes

Data marketplace archetype	Centralized data trading	Centralized data trading with smart contract	De-centralized data trading	Personal data trading
Data Market-place	Quandl	Dawex	IOTA	Datacoup
Value creation	Centralized	Centralized	De-centralized	De-centralized
Value proposition	Anonymized Dynamic datasets	Encrypted Static and dynamic datasets	Encrypted Dynamic datasets	Anonymized Dynamic datasets
Value delivery	API or download Restricted access to data samples B2B No smart contract	API or download Restricted access to data samples B2B Smart contract	API No test data samples B2B Smart contract	Specialized software to access No test data samples C2B Smart contract
Value capture	Freemium pricing Prices set by sellers Fiat currency	Usage based pricing Prices set by sellers Fiat currency	Flat fee pricing Price set by sellers Crypto currency	Usage based pricing Fixed prices Crypto currency

Centralized data trading: With eleven out of twenty data marketplaces, the "standard centralized" archetype was identified as the most common archetype in the dataset. This type of data marketplace has similar characteristics to conventional online marketplaces, offering possibilities to trade data in a simple and efficient manner. Data marketplaces of this archetype do not rely on a specified data origin, data domain, data output type, or pricing model.

Centralized data trading with smart contract: Only one marketplace in the dataset fits this archetype. Although it had a centralized infrastructure, it encrypted the data stored on the marketplace and supported smart contracting with a blockchain. Due to its centralized characteristic, this specific archetype of data marketplace supports straightforward trading of data,

while addressing security and legal issues that can occur in centralized data marketplaces.

Decentralized data trading: Five out of eleven data marketplaces fit into this category. This archetype relies on decentralized infrastructure typical for smart contracting. Marketplaces of this archetype guarantee data quality. They sell self-generated, dynamic data. None of the investigated data marketplaces provided additional services. Four out of five of the data marketplaces following the "standard decentralized" archetype and supported solely cryptocurrency, while one supported both crypto- and fiat currency.

Personal data trading: Three out of twenty marketplaces explicitly allowed users of the data marketplace to expose data for trading. Therefore, this archetype has a consumer-to-business characteristic and operates with user-generated personal data. Data trading is performed through use of simple, specialized software.

5. Discussion and implications

In this research, we used the theoretical lens of business models to propose a taxonomy for data marketplaces. The developed taxonomy consists of dimensions and characteristics derived from conceptual considerations [17, 22, 29, 40, 41, 45, 46] with the addition of new categories, identified by using empirical material from a sample of twenty data marketplaces. Table 4 in the Appendix outlines all conceptual and empirical dimensions and characteristics used for taxonomy building as well as their respective origins. To build the taxonomy, only conceptual characteristics that occurred in the empirical material were subsequently considered. The taxonomy was structured in accordance with the basic business model elements [34, 47], as illustrated in Table 2. Comparing our results with those from previous investigations on the topic [38] indicates that data marketplaces are still evolving, and a dominant business model of data marketplaces is yet to emerge.

Looking at Tables 2 and 4, a main criterion that separated data marketplaces was whether data marketplaces stored their data in a centralized or decentralized manner (e.g., using a blockchain). In that regard, anonymity and encryption of data were identified as major aspects that differentiated data marketplaces. It was surprising that only half of the investigated data marketplaces provided information on this specific topic.

In addition, using the established taxonomy as well as the final sample of twenty data marketplaces, four business model archetypes of data marketplaces were derived: *Centralized data trading, Centralized data trading with smart contract, Decentralized data trading* and *Personal data trading*. See Table 3 for the detailed

characteristics of each archetype. Interestingly, despite their differences, all four archetypes rely on dynamic datasets and privacy measures in their value proposition.

The derived taxonomy as well as the established archetypes provide an overview of the current business models of data marketplaces, subsequently extending the findings of [38]. Our results also allow researchers and practitioners to easily anchor and communicate the dimensions and characteristics of data marketplaces. In addition, the established taxonomy can be used as a basis to design business models of data marketplaces in the future. If required, following the process proposed by [30], the taxonomy can easily be extended to include additional elements. This is especially useful since investigations on business models used in data marketplaces represent a rather new and rapidly evolving research area, where new characteristics or even new dimensions of data marketplaces business models are likely to emerge. Furthermore, as business models represent linkages between strategies and strategy implementation [2] the established business model archetypes contribute to the literature on digital business strategies [4, 21].

6. Limitations and further research

Our research was subject to several limitations. *First*, limiting the investigation to data marketplaces descriptions in the English language that included the keywords "buying" "selling" potentially biased the data selection. This may have contributed to the low number of twenty data marketplaces in the final selection. Further research could be conducted to address this issue by also investigating data marketplaces on a broader basis, for example, by including languages other than English (e.g., Mandarin or Russian) as well as by investigating business models of open data marketplaces. In addition, increasing the number of investigated cases could allow a further evaluation of the established taxonomy as well as a quantitative cluster analysis of data marketplaces. *Second*, although we strictly followed the process of taxonomy building proposed by [30], additional dimensions and characteristics of data marketplaces might be uncovered in future research. This is especially the case since we assume that business models are subject to change, for example, through interactions with their environment [13]. Furthermore, due to the emerging nature of data marketplaces, it is clear that the set of marketplaces used in this analysis only captured a snapshot of what is occurring, rather than allow for long-term observations. It is to be expected that new players will join the market, while others will not be able to cope with challenges, even in the near future. Moreover, investigating the frequency of characteristics of data marketplaces holds

potential for further research, since not all marketplaces offered information on all identified characteristics of the taxonomy at the time of analysis. *Third*, the possibility of misinterpretation of the empirical data cannot be ruled out. However, in order to minimize the probability of misinterpretations, multiple iterations on the collected information on data marketplaces were performed. *Fourth*, this research did not involve direct communication with marketplace providers. However, we assume that including the views of data marketplace providers as part of future research could offer valuable insights into the design of business models for data marketplaces. Future researchers could address this issue by carrying out in-depth case studies of data marketplaces and their respective providers. In addition, such in-depth case studies could include an evaluation of the proposed taxonomy [44]. Our proposed system of dimensions and characteristics as well as the established taxonomy subsequently form a solid basis for future research on business models of data marketplaces.

7. Acknowledgements

8. References

[1] Agarwal, A., M. Dahleh, and T. Sarkar, A Marketplace for Data: An Algorithmic Solution, 2019.

[2] Amit, R. and C. Zott, "Creating Value Through Business Model Innovation", MIT Sloan Management Review, 53(2), 2012, pp. 41–49.

[3] Arrow, K.J., "Economic Welfare and the Allocation of Resources for Invention", in The Rate and Direction of Inventive Activity: Economic and Social Factors. 1962. Princeton University Press.

[4] Bharadwaj, A., O.A. El Sawy, P.A. Pavlou, Venkatraman, and N., "Digital Business Strategy: Toward a Next Generation of Insights", MIS Quarterly, 37(2), 2013, pp. 471–482.

[5] Björkdahl, J., "Technology cross-fertilization and the business model: The case of integrating ICTs in mechanical engineering products", Research Policy, 38(9), 2009, pp. 1468–1477.

[6] Bock, M. and M. Wiener, "Towards a Taxonomy of Digital Business Models - Conceptual Dimensions and Empirical Illustrations", Proceedings of the Thirty-Eighth International Conference on Information Systems (ICIS), Seoul, 2017.

[7] Burkhart, T., J. Krumeich, D. Werth, and P. and Loos, "Analyzing the Business Model Concept â•fl A Comprehensive Classification of Literature", Proceedings of the Thirty Second International Conference on Information Systems (ICIS), Shanghai 2011, 2011.

[8] Carnelley, P., H. Schwenk, G. Cattaneo, G. Micheletti, and D. Osimo, Europe's Data Marketplaces: Current Status and Future Perspectives, 12.06.2016.

[9] Chesbrough, H., "Business Model Innovation: Opportunities and Barriers", Long Range Planning, 43(2-3), 2010, pp. 354–363.

[10] Chesbrough, H. and R.S. Rosenbloom, "The role of the business model in capturing value from innovation: Evidence from Xerox Corporation's technology spin-off companies", Industrial and Corporate Change, 11(3), 2002, pp. 529–555.

[11] Dahlberg, T. and T. Nokkala, "Willingnes to Share Supply Chain Data in an Ecosystem Governed Platform: An Interview Study", in 32nd Bled eConference Humanizing Technology for a Sustainable Society, A. Pucihar, Editor. 2019.

[12] Deichmann, J., K. Heineke, T. Reinbacher, and D. Wee, Creating a successful Internet of Things data marketplace.

[13] Dellyana, D., T.M. Simatupang, and W. Dhewanto, "Managing the Actor's Network, Business Model and Business Model Innovation to Increase Value of the Multidimensional Value Networks", International Journal of Business and Society, 19(1), 2018, pp. 209–218.

[14] Doty, D.H. and W.H. Glick, "Typologies As a Unique Form Of Theory Building: Toward Improved Understanding and Modeling", Academy of Management Review, 19(2), 1994, pp. 230–251.

[15] Eickhoff, M., J. Muntermann, and T. Weinrich, "What do FinTechs actually do?: A Taxonomy of FinTech Business Models", Proceedings of the Thirty-Eighth International Conference on Information Systems (ICIS), Seoul, 2017.

[16] Eisenmann, T., G. Parker, and M.W. van Alstyne, "Strategies for Two-Sided Markets", Harvard Business Review, 2006.

[17] Fricker, S.A. and Y.V. Maksimov, "Pricing of data products in data marketplaces", International Conference of Software Business, 2017, pp. 49–66.

[18] Gawer, A., "Bridging differing perspectives on technological platforms: Toward an integrative framework", Research Policy, 43(7), 2014, pp. 1239–1249.

[19] Günther, W.A., M.H. Rezazade Mehrizi, M. Huysman, and F. Feldberg, "Debating big data: A literature review on realizing value from big data", The Journal of Strategic Information Systems, 26(3), 2017, pp. 191–209.

[20] Hartmann, P.M., M. Zaki, N. Feldmann, and A. Neely, "Capturing value from big data – a taxonomy of data-driven business models used by start-up firms", International Journal of Operations & Production Management, 36(10), 2016, pp. 1382–1406.

[21] Hodapp, D., G. Remane, and A. Hanelt, "Business Models for Internet of Things Platforms: Empirical Development of a Taxonomy and Archetypes", Proceedings of the 14th International Conference on Wirtschaftsinformatik, 2019.

[22] Koutroumpos, P., A. Leiponen, and L.D.W. Thomas, The (Unfulfilled) Potential of Data Marketplaces, ETLA Working Papers No 53, 2017.

[23] Kushal, A., S. Moorthy, and V. Kumar, Pricing for Data Markets, Technical Report, 2012.

[24] Liang, F., W. Yu, D. An, Q. Yang, X. Fu, and W. Zhao, "A survey on big data market: Pricing, trading and protection", IEEE Access, 6, 2018, pp. 15132–15154.

[25] Loebbecke, C. and A. Picot, "Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda", Journal of Strategic Information Systems, 24, 2015, pp. 149–157.

[26] Magretta, J., "Why Business Models Matter", Harvard Business Review, 80(5), 2002, pp. 86–92.

[27] Mamonov, S. and T.M. Triantoro, "The strategic value of data resources in emergent industries", International Journal of Information Management, 39, 2018, pp. 146–155.

[28] Moody, D. and P. Walsh, "Measuring The Value of Information: An Asset Valuation Approach", ECIS 1999 Proceedings (Proceedings of the Ninth European Conference on Information Systems (ECIS), 1999), 1999.

[29] Muschalle, A., F. Stahl, A. Löser, and G. Vossen, "Pricing Approaches for Data Markets", 2012.

[30] Nickerson, R.C., U. Varshney, and J. Muntermann, "A method for taxonomy development and its application in information systems", European Journal of Information Systems, 22(3), 2013, pp. 336–359.

[31] Osterwalder, A. and Y. Pigneur, Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, 1st edn., John Wiley & Sons, 2010.

[32] Otto, B. and S. Aier, "Business Models in the Data Economy: A Case Study from the Business Partner Data Domain", 11th International Conference on Wirtschaftsinformatik, 2013, pp. 475–489.

[33] Özyılmaz, K.R., M. Doğan, and A. Yurdakul, "IDMoB: IoT Data Marketplace on Blockchain", 2018, pp. 11–19.

[34] Remane, G., A. Hanelt, J. Tesch, and L.M. Kolbe, "The business model pattern database: A tool for systematic business model innovation", International Journal of Innovation Management, 21(01), 2017, p. 1750004.

[35] Remane, G., R.C. Nickerson, A. Hanelt, J. Tesch, and L.M. Kolbe, "A-Taxonomy-of-Carsharing-Business-Models", Proceedings of the Thirty-Seventh International Conference on Information Systems (ICIS), Dublin, 2016.

[36] Schmid, B.F. and M.A. Lindemann, "Elements of a reference model for electronic markets", in Proceedings of the Thirty-First Hawaii International Conference on System Sciences, Thirty-First Hawaii International Conference on System Sciences, Kohala Coast, HI, USA, 6-9 Jan. 1998. IEEE Comput. Soc.

[37] Shapiro, C. and H.R. Varian, Information rules: A strategic guide to the network economy, 19th edn., Harvard Business School Press, Boston, 2010.

[38] Spiekermann, M., "Data Marketplaces: Trends and Monetisation of Data Goods", Intereconomics, 54(4), 2019, pp. 208–216.

[39] Spiekermann, M., D. Tebernum, S. Wenzel, and B. Otto, "A Metadata Model for Data Goods", Multikonferenz Wirtschaftsinformatik 2018, 2018, pp. 326–337.

[40] Stahl, F., F. Schomm, L. Vomfell, and G. Vossen, "Marketplaces for Digital Data: Quo Vadis?", Computer and Information Science, 10(4), 2017, pp. 22–37.

[41] Stahl, F., F. Schomm, and G. Vossen, The data marketplace survey revisited, Working Papers, ERCIS - European Research Center for Information Systems, No. 18, 2014.

[42] Stahl, F., F. Schomm, G. Vossen, and L. Vomfell, "A classification framework for data marketplaces", Vietnam Journal of Computer Science, 3(3), 2016, pp. 137–143.

[43] Stuckenberg, S., E. Fielt, and T. Loser, "The Impact Of Software-As-A-Service On Business Models Of Leading Software Vendors. Experiences From Three Exploratory Case Studies", PACIS 2011 Proceedings, 2011.

[44] Szopinski, D., T. Schoormann, and D. Kundisch, "Because your taxonomy is worth it: Towards a framework for taxonomy evaluation", Proceedings of the Twenty-Seventh European Conference on Information Systems (ECIS), Stockholm 2019, 2019.

[45] Täuscher, K., Business Models in the Digital Economy: An Empirical Classification of Digital Marketplaces, 2016.

[46] Täuscher, K. and S.M. Laudien, "Understanding platform business models: A mixed methods study of marketplaces", European Management Journal, 36(3), 2018, pp. 319–329.

[47] Teece, D.J., "Business Models, Business Strategy and Innovation", Long Range Planning, 43(2-3), 2010, pp. 172–194.

[48] Thomas, L.D.W. and A. Leiponen, "Big Data Commercialization", IEEE Engineering Management Review, 44(2), 2016, pp. 74–90.

[49] Veit, D., E. Clemons, A. Benlian, P. Buxmann, T. Hess, D. Kundisch, J.M. Leimeister, and P. Loos, "Business Models: An Information Systems Research Agenda", Business & Information Systems Engineering, 6(1), 2014, pp. 45–53.

[50] Vom Brocke, J., A. Simons, B. Niehaves, B. Niehaves, K. Reimer, R. Plattfaut, and A. Cleven, "Reconstructing the Giant: On the Importance of Rigour in documenting the Literature Search Process", ECIS 2009 Proceedings (Proceedings of the Sixteenths European Conference on Information Systems (ECIS)), 2009.

[51] Wirtz, B.W., A. Pistoia, S. Ullrich, and V. Göttel, "Business Models: Origin, Development and Future Research Perspectives", Long Range Planning, 49(1), 2016, pp. 36–54.

[52] Wixom, B.H., "Cashing in on your data", Center for Information Systems Research, Sloan School of Management, Cambridge, MA: Massachusetts Institute of Technology. Research Briefing, 14(8), 2014.

[53] Wixom, B.H. and J.W. Ross, "How to Monetize Your Data", MIT Sloan Management Review, 58(3), 2017, pp. 10–13.

[54] Woerner, S.L. and B.H. Wixom, "Big data: Extending the business strategy toolbox", Journal of Information Technology, 30(1), 2015, pp. 60–62.

[55] Yin, R.K., Case study research: Design and methods, 4th edn., SAGE, Los Angeles, 2009.

[56] Zott, C. and R. Amit, "The fit between product market strategy and business model: Implications for firm performance", Strategic Management Journal, 29(1), 2008, pp. 1–26.

[57] Zuiderwijk, A., E. Loukis, C. Alexopoulos, M. Janssen, and K. Jeffery, "Elements for the Development of an Open Data Marketplace", Conference for E-Democracy and Open Governement, 2014, pp. 309–322.

Table 4. Description of dimensions and characteristics of data marketplace business models

Dimension	Origin	Characteristic	Description of characteristic
Platform infrastructure	Conceptual [22]	Centralized	Data are stored and accessed from predefined storage spaces
		Decentralized	Data are stored decentralized (e.g., using blockchain)
Data origin	Conceptual [40]	Internet	Data are gathered from online sources (manually or automatically)
		Self-generated	Data are gathered from private sources
		User-generated	Data are collected from user-inputs (e.g., in exchange for using a service)
		Community	Data are collected from marketplaces and crowdsourcing services
		Authority	Data are collected by institutions with a high level of proficiency
Review system	Conceptual [45, 46]	Reviews by users	Reviews directly between buyers and sellers
		Reviews by marketplace	Data marketplace provides reviews
		None	Data marketplace does not provide reviews
Privacy	Empirical	Anonymized	Data marketplace stores anonymized data
		Encrypted	Data marketplace stores encrypted data
		Both	Data marketplace stores anonymized and encrypted data
Data quality guarantee	Empirical	Yes	Data marketplace guarantees quality of purchased data
		No	
Time relevancy	Conceptual [40]	Static	Offered data does not change after its creation
		Dynamic	Regular updates to dataset needed to keep data valid
	Empirical	Both static and dynamic	Offer both static and dynamic datasets
Pre- purchase testability	Conceptual [40]	Complete access	Customers have complete access before paying for data
		Restricted access	Customers can access only part of the data before prior purchase
		None	Customers can not access data before paying for them
Data output type	Conceptual [40]	JSON	Format for semi-structured data
		CSV/XLS	Tabular data
		Report	Visualized data formats (e.g., PDF, DOC, JPEG)
	Empirical	Multiple options	Data marketplace offers multiple options for data output types
Type of access	Conceptual [40]	API	Use of a predefined protocol interface to access data
		Download	Data are accessed through downloadable file
		Specialized Software	Data marketplace requires designated software to handle data
	Empirical	API and Download	Data can be accessed via API as well as via download
Additional purchase support	Empirical	With additional costs	Data marketplace charges for additional services
		Included in price	Data marketplace offers additional services for free
		No	Data marketplace does not provide additional services
Domain	Conceptual [17, 40]	All/Any	Data marketplace not restricted to a certain domain
		Finance/Economy	Economics related data (e.g., stock market data or pricing data)
		Geo	Geographical positions expressed in coordinates
		Address	Lists of customer information (e.g., mail and E-mail addresses)
		Sensor	Data generated by or used for sensors (e.g., IoT data)
	Empirical	Personal	Data related to private information about individuals
Marketplace participants	Conceptual [46]	B2B	Data marketplace operates exclusively in B2B
	Empirical	C2B	Data marketplace operates exclusively in C2B
		Any	Data marketplace not restricted in terms of buyers and sellers
Smart contract with blockchain	Empirical	Yes	Data marketplace offers an option for smart contracting
		No	Data marketplace does not offer an option for smart contracting
Pricing model	Conceptual [29, 41]	Free	Selected datasets are offered for free
		Usage based	Customers pay proportionally for units (e.g., API-calls or time)
		Package pricing	A selected amount of data is offered for a fixed price
		Flat free tariff	Full access to the data marketplace is offered for a recurring fee
		Two-part tariff	Combines a flat fee tariff with additional usage-based pricing
		Freemium	Basic features offered for free, additional features are unlocked for a fee
Price discovery	Conceptual [45, 46]	Fixed prices	Data marketplace has fixed prices
		Set by sellers	Prices are set by sellers
		Set by buyers	Prices are set by buyers
		Auction	Buyers and sellers are bidding against each other
		Negotiation	Buyer and seller agree on an acceptable price for both parties
Payment currency	Empirical	Crypto-currency	Data marketplace handles payment via crypto-currency
		Fiat-currency	Data marketplace handles payment via fiat currency
		Both	Data marketplace handles payment both with fiat currency and crypto-currency