2021

Creating a Taxonomy of Business Models for Data Marketplaces

Montijn van de Ven

Antragama Ewa Abbas

Zenlin Kwee

Mark de Reuver

CREATING A TAXONOMY OF BUSINESS MODELS FOR DATA MARKETPLACES

MONTIJN VAN DE VEN,[1] ANTRAGAMA EWA ABBAS,[2]
ZENLIN KWEE[2] & MARK DE REUVER[2]

[1] Eindhoven University of Technology, Department of Industrial Engineering and Innovation Sciences, The Netherlands; e-mail: m.r.v.d.ven@tue.nl
[2] Delft University of Technology, Faculty of Technology, Policy and Management, The Netherlands; e-mail: a.e.abbas@tudelft.nl, z.roosenboom-kwee@tudelft.nl, g.a.dereuver@tudelft.nl

Abstract Data marketplaces can fulfil a key role in realizing the data economy by enabling the commercial trading of data between organizations. Although data marketplace research is a quickly evolving domain, there is a lack of understanding about data marketplace business models. As data marketplaces are vastly different, a taxonomy of data marketplace business models is developed in this study. A standard taxonomy development method is followed to develop the taxonomy. The final taxonomy comprises of 4 meta-dimensions, 17 business model dimensions and 59 business model characteristics. The taxonomy can be used to classify data marketplace business models and sheds light on how data marketplaces are a unique type of digital platforms. The results of this research provide a basis for theorizing in this rapidly evolving domain that is quickly becoming important.

Keywords: data marketplace, business model, data trading, taxonomy, dimensions

University of Maribor Press

DOI https://doi.org/10.18690/978-961-286-485-9.23
ISBN 978-961-286-485-9

1 Introduction

As an organization may not always possess the required data to carry out or improve their processes and services, they may wish to purchase these data from other organizations. A data marketplace can enable data purchase by providing a digital platform through which individuals and organizations can exchange data (Stahl et al., 2016; Schomm et al., 2013). In contrast to most other platforms, where data is utilized to improve services or manage customer relationships, on data marketplaces data is actually the product itself (Spiekermann et al., 2018).

Despite the potential benefits of data marketplaces, in practice, very little data is shared or traded via these platforms (Koutroumpis et al., 2020). In general, little research has been conducted on data marketplaces (Thomas & Leiponen, 2016) and data marketplace business models in particular (Fruhwirth et al., 2020; Spiekermann, 2019). As a foundation for research on a novel and diverse phenomenon, a first step is developing a taxonomy, because it can be used to classify data marketplace business models (Lambert, 2015). Two taxonomies of data marketplace business models are currently available in the literature, i.e. those proposed by Fruhwirth et al. (2020) and Spiekermann (2019) respectively.

The existing taxonomies (Fruhwirth et al., 2020; Spiekermann, 2019), however, overlook two main areas which this study aims to address. Firstly, the two studies mostly focus on the classification of multilateral data marketplaces, while in practice data trading often happens via bilaterally negotiated contracts (Koutroumpis et al., 2017). Secondly, the studies view data marketplace business models from a single firm perspective. However, data marketplaces take part in a network of stakeholders involving data analysts, application vendors, algorithm developers, data providers, consultants, licensing entities, and platform providers (Muschalle et al., 2012; Thomas & Leiponen, 2016).

To address the above two under-researched areas, this study develops a taxonomy from a multi-stakeholder perspective on business models. We define a business model as the way a network of stakeholders creates and captures value (Bouwman et al., 2008). This multi-stakeholder perspective allows us to understand the business model for the data ecosystem as a whole. Moreover, we define a data marketplace as a digital system where data is traded as an economic good, that connects data

buyers and data sellers, and facilitates data exchange and financial transactions (Koutroumpis et al., 2020; Stahl et al., 2016). In this way, the term data marketplace is broadly interpreted, to go beyond the already studied multilateral data marketplaces.

The remainder of this paper is structured as follows: in Section 2, the taxonomy development process is described. Subsequently, Section 3 presents the developed taxonomy on the basis of the identified business model dimensions. Lastly, Section 4 provides a conclusion of the research and discusses the scientific contribution, practical relevance and limitations of this study.

2 Taxonomy development process

To develop the taxonomy, we follow the taxonomy development method by Nickerson et al (2013). Meta-characteristics of the taxonomy are defined first. Next, the thirteen ending conditions suggested by Nickerson et al. (2013) were employed. After that, multiple iterations are conducted to refine the taxonomy.

2.1 Meta-characteristics

Meta-characteristics function as overarching characteristics of the object of interest (Nickerson et al., 2013). We use the four business model domains of the STOF ontology (i.e. Service, Technology, Organization and Finance domains as in Bouwman et al. (2008)) as the meta-characteristics of the taxonomy, as the STOF approach takes service as a unit of analysis and employs a multi-stakeholder perspective on business models (Bouwman et al., 2008). This perspective is well-suited for data marketplaces because a network of business actors are involved in and around data marketplaces (Muschalle et al., 2012; Thomas & Leiponen, 2016).

2.2 Literature search

We collected dimensions and characteristics from existing literature. A literature search was conducted to discover existing knowledge about the object of interest (Webster & Watson, 2002). Google Scholar was consulted to find relevant academic sources, using the search string "Data marketplaces" AND ("Business models" OR "Digital platform" OR "Digital marketplace" OR "Data trading" OR "Data

economy"). This string resulted in a total of 359 articles. The articles were scanned based on their title, abstract and relevance, which resulted in a preliminary selection of 17 articles. After making this pre-selection of articles, the full text of the articles was read, which resulted in the exclusion of seven articles that did not explicitly discuss dimensions or characteristics of data marketplace business models. Based on cross-reference of the selected articles, we added four additional articles that presented topic-relevant business model taxonomies to the list. The literature review resulted in a final set of 14 articles as presented in Table 1.

Table 1: Overview of classifications relevant to data marketplace business models

Author(s) (Year)	Type	Citations (dated 14 April 2020)
Schomm et al. (2013)	Dimensions of data providers and data marketplaces	73
Stahl et al. (2014a)		14
Stahl et al. (2014b)		16
Stahl et al. (2017)		12
Stahl et al. (2016)	Classification of electronic marketplaces	30
Koutroumpis et al. (2017)	Market designs for data marketplaces	19
Muschalle et al. (2012)	Pricing models for data marketplaces	74
Fricker and Maksimov (2017)		8
Spiekermann (2019)	Taxonomy of data marketplace business models	9
Fruhwirth et al. (2020)		1
Bock and Wiener (2017)	Taxonomy of digital business models	22
Täuscher (2016)	Taxonomy of marketplace business models	6
Täuscher and Laudien (2018)		153
Hartmann et al. (2014)	Taxonomy of data-driven business models	131

2.3 Selection of empirical cases

To account for the practical relevance of the taxonomy, we conducted desk research between May and July 2020 to build a database of empirical cases of data marketplaces. Sixty-five websites of data marketplaces that were mentioned in existing studies of data marketplaces were included in the database (Carnelley et al.,

2016; Koutroumpis et al., 2020, 2017; Prlja, 2019; Spiekermann, 2019; Stahl et al., 2016). The data discovery platform datarade.ai, a website that provides an overview of over 1,800 data providers and 200 data platforms, was consulted. In total, the search in the repository of datarade.ai resulted in the discovery of an additional set of 187 data marketplaces. To complement the database with cases that were not considered in the existing studies or part of the datarade.ai database, we used the search engine Google to further conduct a desk research. The keywords "data marketplace", "data market" and "data trading platform" were applied during the search. From this search, fifteen data marketplaces were added to the database.

Four criteria were applied to the companies that resulted from the desk research to ensure the relevance of the empirical cases. Firstly, data marketplaces that turned out to be shut down, after inspecting the website, were excluded from the database. Secondly, the companies that did not fit this study's definition of a data marketplace were excluded. This implied that data marketplaces that only provided open data, such as governmental organizations and NGOs, were excluded as these platforms adopt non-commercial business models (Carnelley et al., 2016). Thirdly, data marketplaces that did not have an English version of their website were excluded. Lastly, data marketplaces that were still in the construction phase were excluded. The application of these four criteria to the cases led to the exclusion of 89 cases. Therefore, the final database consisted of 178 cases of data marketplaces.

To analyse the business models of existing data marketplaces, a sample was taken from the database of cases. The empiricist philosophy of classification prescribes to build a taxonomy based on the consideration of many characteristics (Lambert, 2015). Therefore, the cases of data marketplaces in the database were first segmented into groups based on the similarity of their characteristics. The website of datarade.ai categorized data marketplaces based on the type of data traded on the platform. This variable was selected as the leading sampling variable to explore the variation between cases in the database. Based on the available information on datarade.ai and an inspection of the case's website, 138 cases could be labelled by type of data traded on the platform. The remaining 40 cases in the database were labelled based on the classification of the cases in the existing studies (Fruhwirth et al., 2020; Spiekermann, 2019) and through the manual inspection of the companies' website.

The segmentation of data marketplaces by type of data traded on the platform reveals that some segments of data marketplaces in the database were overrepresented. This was especially the case for audience data marketplaces, that constituted over 60% of the cases (N=112). Audience data is combined data about a certain target group of customers, which is much sought after by marketeers. To compensate for the overrepresentation, instead of random sampling, a disproportionate stratified sample of N=40 cases was taken from the database (Daniels, 2011). The final sample of 40 data marketplaces consisted of ten data marketplaces on which any type of data is traded (25% of the sample), four financial and alternative data marketplaces (10%), nine audience data marketplaces (22.5%), six sensor and mobility data marketplaces (15%), four geo data marketplaces (10%) and seven health and personal data marketplaces (17.5%) (available here: https://doi.org/10.4121/14679564.v1).

2.4 Design iterations

Our design phase started with a conceptual-to-empirical approach (Nickerson et al., 2013). In these design iterations, the concepts derived from the literature were compared to the sample of empirical cases. Information on the business models of the cases was collected from publicly available sources such as company websites and news articles. The discovered information fragments were coded using the dimensions and characteristics from the literature review as a guideline (See Table 2). After each case, newly identified characteristics were added to the dimensions of the taxonomy. After two conceptual-to-empirical design iterations, two empirical-to-conceptual iterations were conducted, which resulted in the addition of two dimensions to the taxonomy: *enterprise data marketplace* and *data processing and analytics tools*. After every design iteration, the ending conditions were checked. After two conceptual-to-empirical iterations and two empirical-to-conceptual iterations, both the objective and subjective ending conditions were met. Finally, to test the usefulness of the taxonomy, three mini-case studies were conducted on empirical cases of data marketplaces that were not part of the sample, i.e. Wibson, QueXopa and Advaneo respectively. The taxonomy was found to be useful, as the business models of the cases could be classified based on public information about the cases.

Table 2: Coding examples for the value proposition dimension

Characteristic	Case	Quote
Easy data access and/or tooling	Open:Factset Marketplace	*"FactSet creates data and technology solutions for investment professionals around the world, providing instant access to financial data and analytics that investors use to make crucial decisions."*
	Knoema	*"Knoema is a cloud-based data technology platform that makes data accessible and delivers intelligent data tools to enable data access and discovery."*
Secure data sharing	DAWEX	*"With Dawex Global Data Marketplace providers can highlight the value of their data while retaining full control over the distribution and configuration of usage rights."*
	Snowflake	*"Unlike other data marketplaces, Snowflake Data Marketplace leverages Snowflake's Secure Data Sharing technology, which means no data transfer and no need to squeeze data through APIs or use cloud storage."*
High quality and unique data	Amazon DSP	*"Use exclusive Amazon audiences to reach your ideal audience on and off Amazon."*
	Datax	*"Quality business data for better sales leads – Any campaign is only as good as the data it's built on – so make sure yours is the best.*

3 Taxonomy of Data Marketplace Business Models

The final taxonomy consists of 4 meta-dimensions, 17 dimensions and 59 characteristics and is presented in Table 3. In the following sections, the data marketplace business model dimensions are discussed per meta-dimension (STOF).

Table 3: Taxonomy of data marketplace business models

	Dimension	Characteristics			
Service domain	Value proposition	Easy data access and/or tooling	Secure data sharing	High quality and unique data	All services in a single platform
	Enterprise data marketplace	Yes		No	

	Data processing and analytics tools	Yes			No		
	Marketplace participants	B2B		C2B		Any	
	Industry domain	Any data	Geo data	Financial & Alternative data	Health & Personal data	Audience data	Sensor & Mobility data
	Geographic scope	Global		Regional		Local	
	Time frame	Static		Up-to-date	(Near) real-time		Multiple
Technology domain	Platform architecture	Centralized			Decentralized		
	Data access	API		Download	Specialized software		Multiple options
	Data source	Self-generated		Customer provided data	Acquired data		Multiple sources
Organization domain	Matching mechanism	One-to-one		One-to-many	Many-to-one		Many-to-Many
	Platform sponsor	Private		Consortium		Independent	
Finance domain	Revenue model	Commissions		Subscriptions	Usage fees		Asset sales
	Pricing model	Freemium		Pay-per-use	Flat fee tariff	Package based pricing	Multiple
	Price discovery	Set by buyers		Negotiation	Set by marketplace provider		Set by external sellers
	Smart contract	Yes			No		
	Payment currency	Fiat money			Cryptocurrency		

3.1 Service domain

The **value proposition** is a statement that indicates the proposed value that an enterprise intends to deliver to the customer (Bouwman et al., 2008). It often describes how customers can benefit from using the service and how the enterprise aims to set itself apart from the competition. Some data marketplaces offer an **enterprise data marketplace** as an additional service. An enterprise data marketplace functions as a private data marketplace that enables organizations to share data within the company or with external partners, such as suppliers and customers, that are invited by the focal organization. The **data processing and analytics tools** characteristic indicates whether a data marketplace offers additional tooling on top of the data, to perform analytics activities on proprietary data or data bought via the platform. The **marketplace participants** dimension describes the type of participants that are allowed to register and exchange data on the marketplace. While most data marketplaces allow the exchange of any type of data on their marketplace, some data marketplaces focus their data offering towards a specific **industry domain**. The **geographic scope** describes the regions in which the data marketplace is operating and available to users (Täuscher & Laudien, 2018; Täuscher, 2016). The **time frame** dimension describes whether or not the data needs frequent updates to maintain the relevancy of the data (Schomm et al., 2013).

3.2 Technology domain

Data marketplaces may adopt two types of **platform architectures**: centralized or decentralized (Koutroumpis et al., 2017). In the centralized approach, data providers offer their data products via a predefined centralized location on the platform, such as a cloud repository. In decentralized platforms, the data products remain at the data provider and the data is traded using distributed ledger technologies such as blockchain. Platform providers may provide **access to the data** in a number of different ways (Schomm et al, 2013). The **data source** dimension describes the origin where the data was gathered or collected by the data marketplace platform (Hartmann et al., 2014).

3.3 Organization domain

The **matching mechanism** determines the number of parties on each side of the platform (Koutroumpis et al., 2017). Besides multilateral data marketplaces, three more types of data marketplaces exist: bilateral data marketplaces (one-to-one matching), dispersal data marketplaces (one-to-many matching), and harvest data marketplaces (many-to-one matching). The **platform sponsor** can be a private individual or a group, a consortium of buyers or sellers, or an individual or a group that is independent of other market players (Stahl et al., 2017, 2016).

3.4 Finance domain

The **revenue** dimension describes the main source of revenue for a data marketplace (Spiekermann, 2019; Täuscher & Laudien, 2018; Täuscher, 2016). The **pricing model** specifies how the final price for the data good or service is composed (Fruhwirth et al., 2020; Schomm et al., 2013; Spiekermann, 2019; Täuscher & Laudien, 2018; Täuscher, 2016). A **price discovery** function allows buyers and sellers on the marketplace to determine a transaction price which they both agree on (Bakos, 1998). Data marketplaces may implement **smart contracts** to enhance transparency and to enforce trust among marketplace participants (Fruhwirth et al., 2020). The **payment currency** dimension explicates which currencies are accepted for the payments made by marketplace participants (Fruhwirth et al., 2020).

4 Discussion and Conclusion

The developed taxonomy of data marketplace business model has two key scientific contributions. First, the results of the study contribute to the scarce knowledge about data marketplaces and their respective business models (Thomas & Leiponen, 2016). This study adopts a multi-stakeholder perspective on data marketplace business models by emphasizing the roles in the data ecosystem. The taxonomy provides an overview of contemporary knowledge about data marketplace business models and exposes new business model alterations that have emerged in practice. A second contribution made by this study is related to the interpretation of a data marketplace. Existing taxonomies (Fruhwirth et al., 2020; Spiekermann, 2019) focus on studying one type of data marketplaces: multilateral data marketplaces (Koutroumpis et al., 2017). In our study, data marketplaces are more broadly

interpreted as digital systems for trading data as an economic good, that connect buyers and sellers, and facilitate data exchange and financial transactions. This allows us to identify additional business model dimensions, which are not part of existing taxonomies: enterprise data marketplace, data processing and analytics tools, geographic scope, matching mechanism and platform sponsor. By eliciting how data marketplace business models differ, we provide a basis for fine-grained theory development, which is often lacking in platform studies (De Reuver et al., 2018).

The developed taxonomy can guide decision-makers who are exploring the options of setting up a data marketplace or to join an existing data marketplace. An improved understanding about data marketplace business models may help to achieve commercialization, that will make data more accessible and exploitable to individuals, businesses and authorities.

Although we took a systematic approach, subjectivity in assessing the cases may pose a limitation. We dealt with this by conducting multiple iterations and re-interpretations of the data. Further, not all data marketplace companies disclose sufficient information about all of their business model characteristics. Therefore, not all empirical cases could be classified into all of the conceptually derived dimensions. This was especially the case for financially related dimensions such as revenue partners and cost categories (Täuscher & Laudien, 2018). Lastly, as in any taxonomy development study, our study is limited to the current set of phenomena that exist in practice. Hence, future research may update our taxonomy in light of fundamentally new data marketplace types.

Data marketplaces pose a foundation for the data economy: they enable firms to access external data to drive their business and to profit from selling their own data. The EU is investing heavily in data marketplaces in the years to come (European Commission, 2020). At the same time, ambiguity pertains over what constitutes a viable data marketplace business model. Our taxonomy takes a broad and multi-stakeholder perspective to data marketplaces, going beyond the single-firm multilateral perspective of extant taxonomies. We argue that such a broad conceptual basis is needed to advance scholarly understanding of ecosystems in the data economy and to unlock the potential of trading data for a functioning data economy.

Acknowledgments

The research leading to these results has received funding from the European Union's Horizon 2020 Program, under grant agreement 871481 – Trusted Secure Data Sharing Space (TRUSTS).

References

Bakos, Y. (1998). The emerging role of electronic marketplaces on the Internet. Communications of the ACM, 41(8), 35-42.

Bock, M., & Wiener, M. (2017). Towards a Taxonomy of Digital Business Models-Conceptual Dimensions and Empirical Illustrations. In ICIS.

Bouwman, H., de Vos, H., & Haaker, T. (Eds.). (2008). Mobile service innovation and business models. Springer Science & Business Media.

Carnelley, P., Schwenk, H., Cattaneo, G., Micheletti, G., & Osimo, D. (2013). Europe's data marketplaces—current status and future perspectives,'. European Data Market SMART, 63.

Daniel, J. (2011). Sampling essentials: Practical guidelines for making sampling choices. Sage Publications.

de Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: a research agenda. Journal of Information Technology, 33(2), 124-135.

European Commission. (2020). A European Strategy for Data [Press release]. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN

Fricker, S. A., & Maksimov, Y. V. (2017). Pricing of data products in data marketplaces. In International Conference of Software Business (pp. 49-66). Springer, Cham.

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering Business Models of Data Marketplaces. In Proceedings of the 53rd Hawaii International Conference on System Sciences.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). Big data for big business? A taxonomy of data-driven business models used by start-up firms. Cambridge Service Alliance, 1-29.

Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2017). The (unfulfilled) potential of data marketplaces (No. 53). ETLA Working Papers.

Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2020). Markets for data. Industrial and Corporate Change, 29(3), 645-660.

Lambert, S. (2015). The importance of classification to business model research. Journal of Business Models, 3(1).

Muschalle, A., Stahl, F., Löser, A., & Vossen, G. (2012). Pricing approaches for data markets. In International workshop on business intelligence for the real-time enterprise (pp. 129-144). Springer, Berlin, Heidelberg.

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

Prlja, E. (2019). Discovering Business Models of Data Marketplaces [Master's thesis]. Graz University of Technology, Graz, Austria.

Schomm, F., Stahl, F., & Vossen, G. (2013). Marketplaces for data: an initial survey. ACM SIGMOD Record, 42(1), 15-26.

Spiekermann, M., Tebernum, D., Wenzel, S., & Otto, B. (2018). A metadata model for data goods. In Multikonferenz Wirtschaftsinformatik (Vol. 2018, pp. 326-337).

Spiekermann, M. (2019). Data marketplaces: Trends and monetisation of data goods. Intereconomics, 54(4), 208-216.

Stahl, F., Schomm, F., & Vossen, G. (2014a). Data Marketplaces: An Emerging Species. In DB&IS (pp. 145-158).

Stahl, F., Schomm, F., & Vossen, G. (2014b). The data marketplace survey revisited (No. 18). ERCIS Working Paper.

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. Vietnam Journal of Computer Science, 3(3), 137-143.

Stahl, F., Schomm, F., Vomfell, L., & Vossen, G. (2017). Marketplaces for Digital Data: Quo Vadis?. Computer and Information Science, 10(4).

Täuscher, K. (2016). Business models in the digital economy: an empirical study of digital marketplaces. Fraunhofer MOEZ. Fraunhofer Center for International Management and Knowledge Economy, Städtisches Kaufhaus Leipzig, Neumarkt, 9-19.

Täuscher, K., & Laudien, S. M. (2018). Understanding platform business models: A mixed methods study of marketplaces. European Management Journal, 36(3), 319-329.

Thomas, L. D., & Leiponen, A. (2016). Big data commercialization. IEEE Engineering Management Review, 44(2), 74-90.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, xiii-xxiii