A Business Model Taxonomy for Data Marketplaces

Data Trade in Various Trading Structures

Rômy Bergman

Master Thesis Report

TU Delft

A Business Model Taxonomy for Data Marketplaces

Data Trade in Various Trading Structures

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Complex Systems Engineering & Management**

Faculty of Technology, Policy and Management

by

Rômy Bergman

Student number: 4369408

To be defended in public on July 30th 2020

Graduation committee

First Supervisor	: Dr. ir., G.A., de Reuver, ICT
Second Supervisor	: Dr. rer. pol., C., Werker, ETI
External Supervisor	: S. Jung, Bosch IoT Lab

Preface

Dear reader,

In front of you lies my master thesis "A Business Model Taxonomy for Data Marketplaces". This thesis is the final work of my Master Complex Systems Engineering and Management at the TU Delft. The concept of data marketplaces fascinated me when the topic was introduced in one of my lectures over a year ago. At the time, I thought I had a good idea of what a data marketplace is and how it could develop. During the past 6 months, I learned more about data marketplaces and their business models from practitioners and researchers. The more I learned, the less I knew. My perception of what a data marketplace is and what it could be changed quite a bit since I started my thesis. I hope that my results contribute to further evolve the concept of data marketplaces and their business models.

There are numerous people who guided me throughout the process whom I would like to express my gratitude to. First of all, Mark de Reuver thank you for your supervision and feedback throughout my research process. Even before I started my thesis you helped me by establishing a connection to the Bosch IoT Lab in St Gallen. I highly appreciate this effort, as it was my aspiration to go there. Furthermore, I enjoyed our Skype calls and always left our meetings with a better sense of direction for my research. Claudia Werker, thank you for your invaluable feedback and for challenging me to step outside of my comfort zone. Although this came with some resistance from my side in the beginning, I am glad that you pushed me to start my research with an explorative approach. It taught me to be more open minded in my approach and gain new insights from practitioners early on in my research. Next, I would like to thank Sven Jung from the Bosch IoT Lab for the warm welcome in St Gallen. Although my time was cut short in Switzerland, you continued to guide me from a distance. Our discussions about business models and data marketplaces helped me to improve my storyline. Further, I would like to thank the interviewees who were willing to answer my questions and share their expertise about data marketplaces.

A special thanks to my parents and sister for their continuous support. They saw me struggle at times and never failed to motivate me. The words of encouragement from my parents and occasional treats fueled me to study until late hours. The external view on my thesis from my sister helped me improve my line of argumentation. I could not have done it without you!

Best regards,

Rômy Bergman

July 20th, 2020

Executive Summary

Companies are increasingly reliant on internal and external data sources to innovate their products and services. Currently, companies produce data for their own usage and store it in data silos afterwards. This hinders secondary data usage, when companies reuse external data. Data marketplaces are multi-sided platforms where data sellers, buyers and third-party service providers can trade data. However, companies rarely trade industrial datasets on multilateral data marketplaces (Koutroumpis, Leiponen & Thomas, 2017). These data marketplaces are difficult to set up and remain conceptual.

Data marketplace owners apply business models to transform technical ideas into functioning value propositions. Research about business models of data marketplaces is required to further advance the development of data marketplaces. However, business model literature of data marketplaces is limited. Researchers mainly focus on theoretical concepts that are not implemented in practice (Constantinides, Henfridsson & Parker, 2018). Thus, it remains unclear what business models data marketplace owners apply in practice.

Researchers apply taxonomies to show the essential elements of objects and to compare cases. They can use taxonomies to clarify what business models data marketplace owners apply. However, few business model taxonomies for data marketplaces exist. Taxonomies were created by Spiekermann (2019) and Fruhwirth, Rachinger & Prlja (2020), but they lack on the following fronts:

- Their taxonomies vary in the included business model dimensions. This indicates misalignment or misinterpretation of the dimensions.
- Their taxonomies are based on cross-industry data marketplaces. This leads to a general interpretation of business model dimensions and characteristics.
- Their taxonomies are based on multilateral data marketplaces. However, these data marketplaces remain conceptual ideas without a viable business model. Data marketplaces with a hierarchical orientation and private ownership are not considered in these taxonomies while in practice these marketplaces have established business models.

Our objective is to clarify what business models data marketplace owners apply in the business-to-business (B2B) automotive industry. We focus on the B2B automotive industry because this industry has established data marketplaces as identified by Martens & Mueller-langer (2018). Through investigating this specific industry, business model components could be identified that data marketplace owners apply successfully in practice. To achieve our objective, we designed a business model taxonomy for data marketplaces and subsequently derived business model archetypes from our taxonomy. We bridge the previously mentioned gaps as follows:

- Our taxonomy includes business model dimensions that we derived from interviews with data marketplace owners. We aligned these dimensions with the dimensions from the taxonomies developed thus far, thereby decreasing variability among taxonomies by Spiekermann (2019) and Fruhwirth et al. (2020).
- We classified data marketplaces from the B2B automotive industry in our taxonomy to be specific in our interpretation of business model dimensions and characteristics.
- We included different types of data marketplaces that vary in their orientation and ownership. The orientation of a data marketplace refers to the coordination of data trade in a hierarchical or market structure. Ownership indicates whether one private company, a number of companies or an independent party owns the data marketplace. In practice, data marketplaces with a hierarchical orientation and private ownership mainly occur. Therefore, we included these data marketplace types, which sets our taxonomy apart from the ones created by Spiekermann (2019) and Fruhwirth et al. (2020).

The following research question is addressed in this thesis: *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?*

To design our taxonomy, we followed the taxonomy development approach by Nickerson, Varshney & Muntermann (2013). They suggest to iteratively induce and deduce dimensions and characteristics to create a taxonomy. Because they do not explicitly state the methods to induce or deduce concepts, we extended their approach with other research methods.

First, we induced business model dimensions using the Grounded Theory Method. Grounded theory is constructed through inductive reasoning, starting with information gathered from interviews, reports and other data materials. This formed the first explorative step in designing our taxonomy as is advised by Nickerson et al. (2013) in areas where little data about the research domain is available. Based on interviews with seven data marketplace owners, we derived five business model dimensions; contracts, platform infrastructure, data processing activities, revenue streams and pricing mechanisms. Data marketplace owners apply negotiated or standardized *contracts* to incorporate data regulation into their business model and create customer relationships. The dimension contract is applied by data marketplace owners to create value. The centralized or decentralized *platform infrastructure* enables data marketplace owners to perform *data processing activities*. These are key resources and activities of data marketplace owners to deliver value. To capture value from data trade, data marketplace owners receive *revenue streams* from their participants and apply *data pricing mechanisms* to monetize data. Data marketplace owners experience that data pricing is challenging for them and their customers because people are not used to monetize data. Fixed pricing mechanisms are often established by the data sellers instead of dynamic pricing mechanisms that fluctuate based on supply and demand in the market. We aligned the induced dimensions with dimensions that are deduced from the taxonomies by Spiekermann (2019) and Fruhwirth et al. (2020) to compose our preliminary taxonomy.

Next, we induced business model characteristics from a selection of data marketplace cases by applying content analysis. Three data marketplace types are represented in our case selection of six data marketplaces from the B2B automotive industry. First, TomTom and INRIX are data marketplaces with a hierarchical orientation and private ownership. Second, HERE and Caruso are data marketplaces with characteristics from both a hierarchical and market orientation and have consortium ownership. Third, IOTA and Ocean Protocol have a market orientation and independent ownership. The classification of these data marketplaces results in a refined taxonomy of thirteen dimensions and thirty-five characteristics. With these characteristics we can distinguish how data marketplace owners create, deliver and capture value.

Based on a cross-case analysis of the data marketplaces in our taxonomy, we recognized patterns and derived four business model archetypes. These are (i) the *aggregating data marketplace* archetype, applied by TomTom and INRIX, (ii) the *aggregating data marketplace with an additional brokering service* archetype, applied by HERE, (iii) the *consulting data marketplace* archetype, applied by Caruso and (iv) the *facilitating data marketplace* archetype which IOTA and Ocean Protocol apply. Our archetypes show that data marketplace owners create value for their customers in various manners. The data marketplace owners who apply the *aggregating data marketplace*, *aggregating data marketplace with an additional brokering service* or *consulting data marketplace* archetype create value by performing additional services such as a customized map service, reviewing the data quality or offering personal assistance through bilaterally negotiated contracts. By performing these value-adding services, data marketplace owners attract customers. The data marketplace owners who apply the *facilitating data marketplace* archetype perform none of these value-adding services. They aim to increase data accessibility for their customers by focusing on their data brokering service. Data is traded between data sellers and buyers with minimal interference of the data marketplace owner. However, few participants are active on these platforms. With few or no data sellers at their platform, the data marketplace owners cannot fulfill their promise of increased data access. The data marketplace owners who apply the *facilitating data marketplace* aim for dynamic pricing mechanisms but apply fixed pricing mechanisms in reality. Without dynamic pricing mechanisms, the desired competitive environment remains absent. With a lack of increased data accessibility and competitive pricing, these data marketplace owners fail to attract customers.

We contribute to academic knowledge by including data marketplace types ranging from hierarchical to market orientation and private to independent ownership. The inclusion of these data marketplace types enabled the identification of business models that data marketplace owners actually apply in practice. Our business model taxonomy and archetypes serve as an overview to further advance the development of data marketplaces. Data marketplace owners are advised to attract customers by performing value-adding services. They can use our taxonomy to make design choices for their own business model. Our taxonomy may also serve as a tool for practitioners to perform a competitor analysis. Whether practitioners would truly use our taxonomy for such purposes is not evaluated. This is recommended for future research. Furthermore, researchers may classify additional data marketplaces from other industries in our taxonomy to make it more reliable. This may validate our results or generate new business model dimensions and characteristics useful for researchers as well as data marketplace owners.

Contents

List of Tables

List of Figures

List of Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
AV	Autonomous Vehicle
B2B	business-to-business
C2B	consumer-to-business
C2C	consumer-to-consumer
DLT	Distributed Ledger Technology
DM	Data marketplace
EC	European Commission
EV	Electrical Vehicle
GDPR	General Data Protection Regulation
IoT	Internet of Things
NTL	Non-Technical Literature
P2P	peer-to-peer
OEM	Original Equipment Manufacturer
RTTI	Real-Time Traffic Information
SDK	Software Development Kit
SME	Small and Medium-sized Enterprises
TL	Technical Literature

1. Introduction

1.1 Problem Identification

Data is generated everywhere and its value for enterprises is progressively recognized. This is a result of the evolvement of the Internet of Things (IoT) which connects devices, systems and people (Cheng, Longo, Cirillo, Bauer & Kovacs, 2015). The emergence of IoT causes a shift from stocks of data into a constant stream of data (Tiwana, 2013). Enterprises become increasingly reliant on data streams as a resource to further advance their businesses (Hartmann, Zaki, Feldmann & Neely, 2014). Companies use internal and external data streams to improve their processes and innovate existing and novel products or services (Agahari, de Reuver & Fiebig, 2019). However, data is often only utilized for primary usage and stored in data silos afterwards (Perera et al., 2017). According to Thomas & Leiponen (2016) the value of data is in its secondary use, when data from organizations is reused externally.

The potential of secondary data use is targeted by data marketplace owners. In theory, data marketplaces are multi-sided platforms where data marketplace owners, data sellers, data buyers and third-party service providers easily trade, store and access data (Mišura & Žagar, 2016). However in reality, companies rarely trade industrial datasets on multilateral data marketplaces and preferably trade data on bilateral basis (Koutroumpis, Leiponen & Thomas, 2017). Several issues that interfere with the advancement of trade via multilateral data marketplaces have been identified such as data security, user privacy (Park, Youn, Kim, Rhee & Shin, 2018; Spiekermann, 2019), data quality preservation (Koutroumpis et al., 2017; Perera et al., 2017), data monetization and revenue optimization (Mao, Zheng, & Wu, 2019; Spiekermann, 2019).

Business models contribute to solving these issues. Business model frameworks help researchers and practitioners to understand, analyze and communicate strategic design choices as well as inform the design of information systems (Al-Debei & Avison, 2010). Data marketplace owners apply business models to transform technical ideas into functioning value propositions (Amit & Zott, 2001). A business model demonstrates how companies create, deliver and capture value (Teece, 2010). Literature about business models for data marketplaces is fragmented and is still evolving (Fruhwirth, Rachinger & Prlja, 2020). Various researchers discuss individual components of data marketplaces and propose pricing, quality and privacy mechanisms to improve data marketplaces (Mao et al., 2019; Park et al., 2018; Perera et al., 2017). These proposed mechanisms remain theoretical ideas and are not all implemented in practice (Constantinides, Henfridsson & Parker, 2018). Thus, business model literature about data marketplaces comprehends predominantly theoretical components rather than practically applicable components. Therefore, it remains unclear what business models data marketplace owners actually apply in practice.

Taxonomies are suitable artefacts to provide such insights. As stated by Nickerson, Varshney & Muntermann (2013), taxonomies aid researchers and practitioners in deciding on the uniqueness of existing applications or in pointing out possibilities for new developments. Researchers classify objects in taxonomies to reduce complexity and identify similarities and differences between objects (Nickerson et al., 2013). Fruhwirth et al. (2020) and Spiekermann (2019) started with structuring the business models of data marketplaces into taxonomies. They highlight the need to extend the taxonomies with new characteristics due to the fast change data marketplaces undergo.

In this MSc thesis, we research the business models of various data marketplaces. In the remaining paragraphs of this chapter we introduce key components for this research. The state of the art in data marketplaces and their business models are discussed in section 1.2. Next, the scope of the research is specified in section 1.3. Subsequently, the research gap, objective and question are outlined in section 1.4. The research approach is explained in section 1.5. Moreover, the relevance of this thesis to the MSc program Complex Systems Engineering and Management at the TU Delft is elaborated in section 1.6. The chapter is concluded with a reading guide for the following chapters in section 1.7.

1.2 State of the Art

1.2.1 Defining Data Marketplaces

The concept of a data marketplace is fairly new and the definition of a data marketplace is still evolving. Terms such as data intermediaries (Flipsen, 2019), data collaboratives (Susha et al., 2017; van den Broek & van Veenstra, 2015) and data marketplaces (Agahari et al., 2019) can be used interchangeably. We adopt the term "data marketplace" as it stresses the goal to trade data.

To comprehend what a data marketplace entails, the terms *data* and *marketplace* should be clarified. *Data* is a non-physical good. It is the core product that is traded at a data marketplace. Data is an intermediate good, which business owners use to create new products (Koutroumpis et al., 2017). Data can appear in different forms (raw and aggregated). Stahl et al. (2016) require data marketplaces to contain machine-readable data, such as RDF or XML, which we adopt in our definition of a data marketplace. Platforms such as Wikipedia, where data is traded in textual form, are excluded from our definition. *Marketplaces* are the online or offline infrastructures where participants exchange goods (Stahl et al., 2016). There are three main functions a marketplace should fulfill (Bakos, 1998):

1. Match buyers and sellers: the buyer's demand and seller's supply should be matched by determining the product offerings, searching for buyers and sellers and determining the price.
2. Facilitate transactions: mechanisms for logistics and settlement should lead to the transportation of the sold product and transfer of payment.
3. Provide an institutional infrastructure: markets should have mechanisms to enforce laws, rules and regulations to coordinate transactions.

With the previous notions of *data* and *marketplaces*, we create the following definition of a data marketplace: a *data marketplace* <u>matches</u> buyers and sellers, <u>facilitates</u> transactions and <u>provides</u> an institutional infrastructure to trade <u>machine-readable data</u>.

In literature, researchers characterize data marketplaces by the participants who are active on the platform. Four key players are mentioned frequently. These are the data marketplace owner, data sellers, data buyers and third party service providers (Fruhwirth et al., 2020; Koutroumpis et al., 2017; Muschalle et al., 2012; Spiekermann, 2019). Spiekermann (2019) describes the relationships between the four key players. The data marketplace owner hosts the data on the platform. The data is made available by the data seller, who owns the data. Data sellers may be commercial or non-commercial parties. Data is sold to data buyers who are consumers or businesses. Third party service providers leverage datasets and add value to the data. They retrieve data from the data marketplace and upload a transformed dataset. This results in a data marketplace overview as shown in Figure 1.

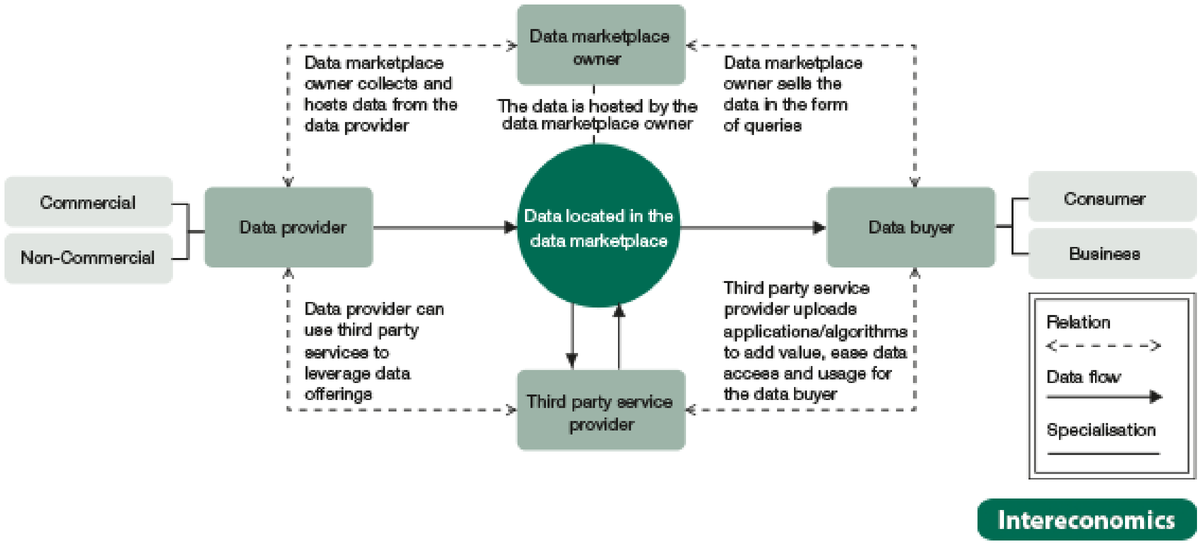

Figure 1: data marketplace overview (Spiekermann, 2019)

According to Spiekermann (2019), data marketplace participants trade data on multilateral basis. As explained in section 1.1, multilateral data marketplaces remain theoretical. In reality data marketplace participants rather trade data on bilateral basis. To research business models of data marketplaces that go beyond theoretical concepts, we include business models of data marketplaces that occur in practice. This indicates that we need to consider different types of data marketplaces.

Stahl et al. (2016) propose a framework that enables classification of data marketplaces in different types. They make use of two determinants: orientation and ownership. Orientation refers to whether the data marketplace owner coordinates data trade in a hierarchical or market trading structure. In data marketplaces with a hierarchical orientation, the data marketplace owner determines the data price and what buyers and sellers are allowed on the platform. In data marketplaces with a market orientation, prices are determined by the buyers and sellers depending on competitive offerings. Ownership indicates whether one private company, a number of companies or an independent party owns the data marketplace. Koutroumpis et al. (2017) maintain a similar classification in which they sort data marketplaces based on their matching mechanism. They distinguish between four types of data marketplaces; one-to-one, one-to-many, many-to-one and many-to-many data marketplaces. First, one-to-one data marketplaces are bilateral marketplaces where two parties are directly connected. One seller will trade with one buyer. Second, at one-to-many data marketplaces there is one seller who trades with many buyers for the same data. In this case, standardized terms of exchange through APIs are maintained, because it is too costly to negotiate data individually. Third, many-to-one data marketplaces allow multiple sellers and one buyer. The sellers usually make their data available to one service provider and receive a service in return for free, as practiced on social media platforms. Fourth, many-to-many data marketplaces are multilateral marketplaces where many sellers and buyers trade data. There is often no specific ownership over the data, but transactions to acquire data are facilitated.

We combine the classifications of Koutroumpis et al. (2017) and Stahl et al. (2016) in Figure 2. This shows the spectrum in which different types of data marketplaces can be classified, depending on their orientation and ownership. Stahl et al. (2016) identify 6 types of marketplaces of which 3 overlap with the one-to-many, many-to-one and many-to-many data marketplaces as defined by Koutroumpis et al. (2017). The one-to-one data marketplace defined by Koutroumpis et al. (2017) is added to the selection, resulting in a total of 7 types.

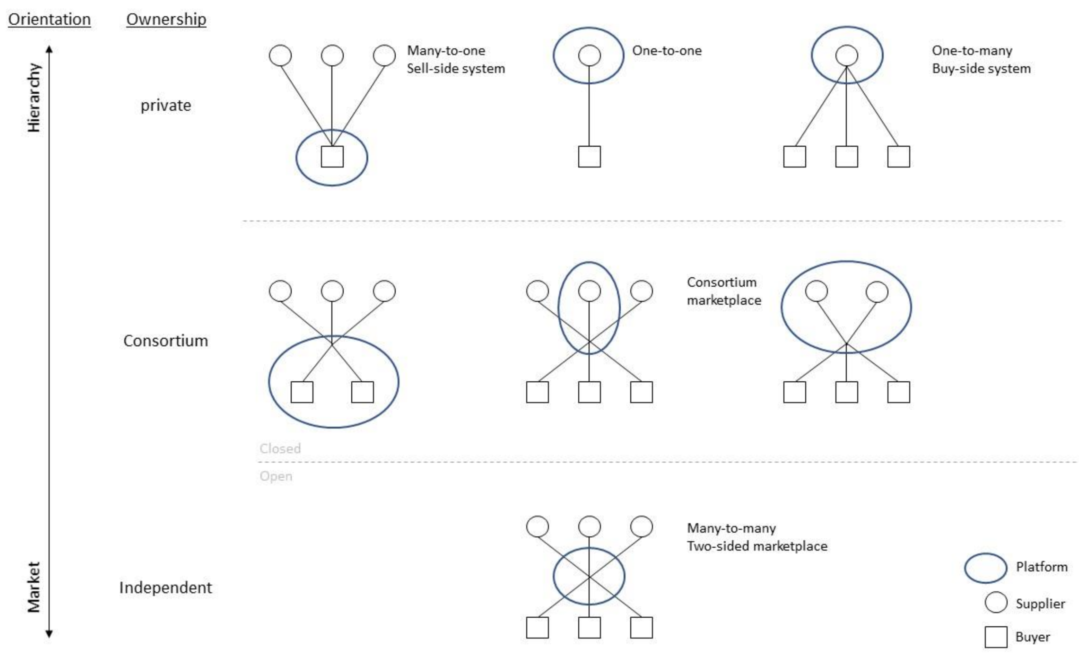

Figure 2: Data marketplace types adapted from Koutroumpis et al. (2017) and Stahl et al (2016)

In this research, data marketplace types differ in terms of hierarchical or market orientation and private, consortium or independent ownership. This expands the definition of a data marketplace by Spiekermann (2019) and allows us to include data marketplaces that range from hierarchical to market orientation. We extend our previous definition of a data marketplace to: a *data marketplace* has a hierarchical or market <u>orientation</u> and private, consortium or independent <u>ownership</u> and <u>matches</u> buyers and sellers, <u>facilitates</u> transactions and <u>provides</u> an institutional infrastructure to trade <u>machine-readable data</u> (see Figure 3).

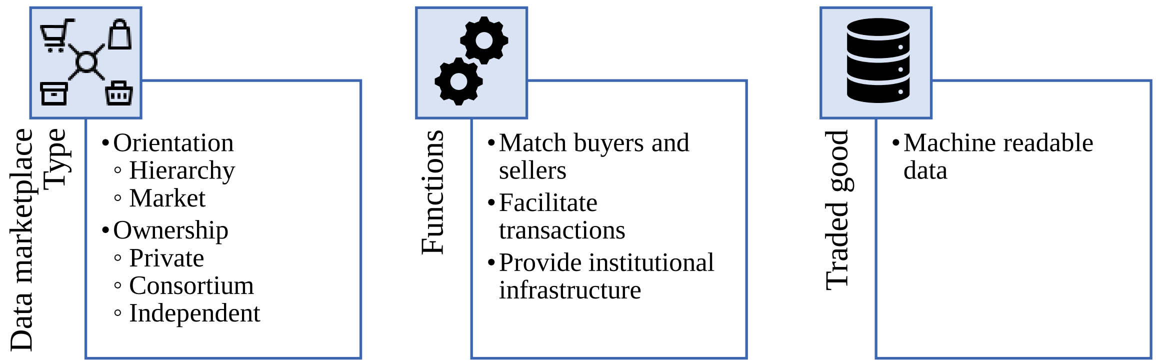

Figure 3: Data marketplace requirements

1.2.2 Business Model Taxonomies for Data Marketplaces

In the emerging research field of data marketplaces, taxonomies are useful to analyze the business models of data marketplaces. Taxonomies are suitable artefacts to analyze and understand a domain (Nickerson et al., 2013; Szopinski et al., 2019). However, few taxonomies of business models for data marketplaces exist. Current research contains two taxonomies that have been developed by Spiekermann (2019) and Fruhwirth et al. (2020).

Spiekermann (2019) identifies value proposition, market positioning, market access, integration, data transformation, architecture, price model and revenue model as business model dimensions. Value adding services, such as data analytics, appeal to data sellers and buyers and are a key success factor in the business model. According to Spiekermann (2019), the acceptance of data exchange at data marketplaces is growing among data sellers and buyers. He identifies a shift towards commercial exchange of data. However, the taxonomy developed by Spiekermann (2019) remains high-level and can be extended with more granular business model characteristics.

Fruhwirth et al. (2020) structure the business model dimensions as value proposition, creation, delivery and capture in their taxonomy. They were able to identify four data marketplace archetypes. These are *centralized data trading, centralized data trading with smart contract*, *decentralized data trading* and *personal data trading*. The archetypes differ in platform infrastructure, privacy and access type. In contrast to Spiekermann (2019), Fruhwirth et al. (2020) do not take market positioning of the platform owner and data transformation activities into consideration. At the same time, Spiekermann (2019) does not consider the dimensions time relevancy and payment currency, which Fruhwirth (2020) does consider. The contrasting qualities indicate misalignment in the included business model dimensions and asks for a revision of their taxonomies.

Both taxonomies are based on data marketplaces with a market orientation. However, these types of data marketplaces appear to be challenging in their set-up and many initiatives fail (Koutroumpis et al., 2017). Spiekermann (2019) attributes the failure of multilateral data marketplaces to the fear data sellers have of losing control over their data, the unwillingness of customers to pay the price for the data and the lack of legal frameworks. Several researchers have developed mechanisms to solve issues that multilateral data marketplace owners face (Mao et al., 2019; Park et al., 2018; Perera et al., 2017). However, the proposed mechanisms remain conceptual (Constantinides et al., 2018). Hence, the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) and the suggested mechanisms in data marketplace literature do not offer insight in business models that data marketplace owners apply in practice.

1.3 Scope

A wide range of participants can trade data on data marketplaces. There are business-to-business (B2B), consumer-to-business (C2B) and peer-to-peer (P2P), also known as consumer-to-consumer (C2C), data marketplaces (Fruhwirth et al., 2020). Some data marketplaces are specialized in one sector while others offer data products in numerous industries that range from transport to retail and from energy to agriculture data. Because the data that is rarely traded concerns industrial datasets that is mainly traded between businesses, we focus on B2B data marketplaces. Our assumption is that industrial datasets are not traded between end-consumers.

Spiekermann (2019) and Fruhwirth (2020) designed business model taxonomies focused on multilateral data marketplaces that trade data across industries. Reoccurring industries that are represented at these marketplaces are the automotive, energy and health industry. Some examples of cross-industry data marketplaces are IOTA, Ocean Protocol and Databroker Dao. However, these data marketplaces are in the beta or proof-of-concept phase. Data marketplaces that are past the conceptual stage and are active in the market are within-industry data marketplaces (Koutroumpis et al., 2017). Koutroumpis et al. (2017) explain that these data marketplaces pool data among participants in one industry. We expect that a business model taxonomy based on within-industry data marketplaces goes beyond theoretical concepts and shows business model characteristics that contribute to the development of data marketplaces in practice. We focus on the B2B automotive industry because this industry has established data marketplaces as identified by Martens & Mueller-langer (2018). They investigated multiple initiatives such as infotainment service platforms and data marketplaces where mobility data is traded between businesses.

Critique to focusing on the automotive industry would be that results are not generalizable for other sectors. However, Spiekermann (2019) and Fruhwirth (2020) already designed taxonomies based on multiple industries. We focus explicitly on one industry to create business model insights in established data marketplaces. Within the timeframe of this master thesis we can analyze data marketplaces from one industry on a deeper level. Future researchers may focus on other industries to combine the results and reach a more extensive taxonomy that is applicable to all industries.

1.4 Research Gap, Objective and Question

As explained in section 1.1, data marketplace owners create value by enabling data trade between multiple parties, but in reality industrial datasets are rarely traded on multilateral basis. Most data marketplaces with a market orientation remain conceptual. The multilateral trade of data at a multi-sided platform, as shown in Figure 1, fits the ideal form of a marketplace. In practice, various types of data marketplaces emerge ranging from hierarchical to market orientation and private to independent ownership, as shown in Figure 2. The question remains how various data marketplaces exactly differ and what components contribute to a data marketplace in practice. Business models can offer insight into these questions as data marketplace owners use them to transform concepts into functioning value propositions (Amit & Zott, 2001). However, literature about business models of data marketplaces is fragmented (Fruhwirth et al., 2020). To create an overview of the business models that data marketplace owners apply, Fruhwirth et al. (2020) and Spiekermann (2019) designed taxonomies. However, there are multiple factors that cause these taxonomies to be incomplete:

- The taxonomies vary in the included business model dimensions. This indicates misalignment or misinterpretation of the dimensions. As explained in section 1.2.1 Spiekermann (2019) includes dimensions such as market positioning and transformation activities, which are excluded by Fruhwirth et al. (2020). On the other hand the taxonomy of Fruhwirth et al. (2020) contains the dimensions time relevancy and payment currency which are not considered by Spiekermann (2019). An explanation why certain dimensions are included or excluded from their taxonomies lacks.
- The taxonomies are based on cross-industry data marketplaces. This leads to a general interpretation of business model dimensions and characteristics.
- The taxonomies are based on multilateral data marketplaces. However, these data marketplaces remain conceptual ideas without a viable business model. Data marketplaces with a hierarchical orientation and private ownership are not considered in these taxonomies while in practice these data marketplaces have established business models.

Our objective is to clarify what business models data marketplace owners apply in the B2B automotive industry. We focus on the B2B automotive industry because this industry has established data marketplaces as identified by Martens & Mueller-langer (2018). Through investigating this specific industry, business model components could be identified that data marketplace owners apply successfully in practice. To achieve our objective, we first design a business model taxonomy for data marketplaces. We bridge the previously mentioned gaps as follows:

- Our taxonomy includes business model dimensions that we derived from interviews with data marketplace owners. We aligned these dimensions with the dimensions from the taxonomies developed thus far, thereby decreasing variability among taxonomies by Spiekermann (2019) and Fruhwirth et al. (2020).
- We classify data marketplaces from the B2B automotive industry in our taxonomy to be specific in our interpretation of business model dimensions and characteristics.
- We include different types of data marketplaces that vary in their orientation and ownership. The orientation of a data marketplace refers to the coordination of data trade in a hierarchical or market structure. Ownership indicates whether one private company, a number of companies or an independent party owns the data marketplace. In practice, data marketplaces with a hierarchical orientation and private ownership mainly occur. Therefore, we include these data

marketplace types, which sets our taxonomy apart from the ones created by Spiekermann (2019) and Fruhwirth et al. (2020).

Second, business model patterns are derived from our taxonomy to generate business model archetypes for data marketplaces. Archetypes comprise the characteristics and dimensions of similar cases (Oberländer et al., 2019). Fruhwirth et al. (2020) generated archetypes, but there is a need to improve these archetypes because there are additional business model dimensions that need to be considered. As acknowledged by Fruhwirth et al. (2020) and Spiekermann (2019), characteristics of data marketplaces and their business models evolve quickly and there is a need for constant extension and adaptation of data marketplace taxonomies and archetypes.

The business model archetypes are linked to the types of data marketplaces. In this thesis, the types of data marketplaces range from hierarchical to market orientation and private to independent ownership. This includes data marketplaces that are past the conceptual stage. Therefore, we contribute to existing scientific literature by identifying business models that data marketplace owners apply in practice. Our main research question is: *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?*

1.5 Research Approach

To identify the business models that data marketplace owners apply, we perform a qualitative research study. We follow the iterative taxonomy development approach by Nickerson et al. (2013) who combine inductive and deductive research (see Figure 4). Their approach offers a systematic way to develop a taxonomy and is widely accepted in the field of information systems (Szopinski, Schoormann & Kundisch, 2019). In the empirical-to-conceptual step, concepts from existing objects are induced. In the conceptual-to-empirical step, concepts are deduced from literature. The combination of both approaches leads to the design of our taxonomy.

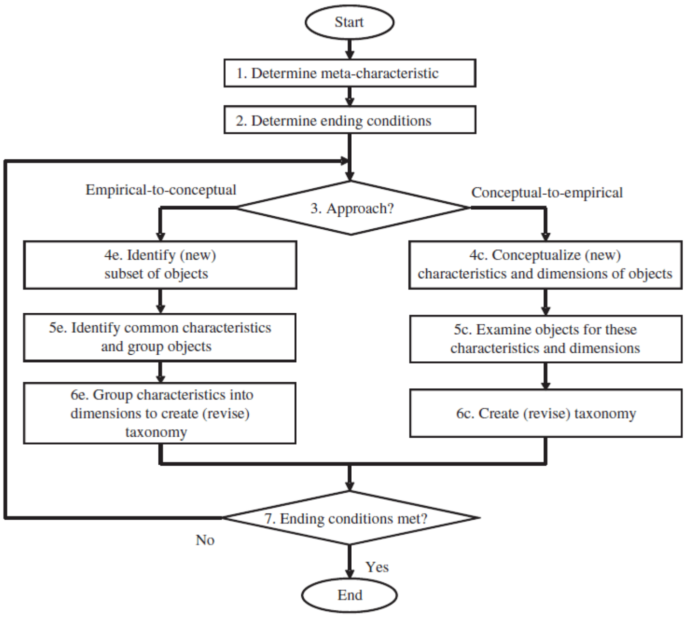

Figure 4: Taxonomy development approach by Nickerson et al. (2013)

The first two steps in the taxonomy development approach are to determine the meta-characteristics and ending conditions (see Figure 4). Nickerson et al. (2013) define meta-characteristics as "the most comprehensive characteristic that will serve as the basis for the choice of characteristics in the taxonomy" (p. 343). The choice in meta-characteristics should be based on the purpose of the taxonomy (Nickerson et al., 2013). In section 1.1, we explained that it is unclear what business models data marketplace owners apply. The purpose of our taxonomy is to identify the business model dimensions and characteristics that data marketplace owners in the B2B automotive industry apply. Thus, the meta-characteristics of our taxonomy need to represent the main business model components. Teece (2010) describes value creation, delivery and capture as the main business model mechanisms. These mechanisms are the chosen meta-characteristics for our taxonomy. In addition, ending conditions are defined to determine when the taxonomy development process is completed (Nickerson et al., 2013). We adopt the objective and subjective ending conditions of Nickerson et al. (2013), as done by Fruhwirth et al. (2020) and Möller et al. (2019). The ending conditions are listed in Table 1.

Table 1: Ending conditions adopted from Nickerson et al. (2013, p. 344)

Objective Ending Conditions	
OE1	All objects or a representative sample of objects have been examined
OE2	No object was merged with a similar object or split into multiple objects in the last iteration
OE3	At least one object is classified under every characteristic of every dimension
OE4	No new dimensions or characteristics were added in the last iteration
OE5	No dimensions or characteristics were merged or split in the last iteration
OE6	Every dimension is unique and not repeated (i.e., there is no dimension duplication)
OE7	Every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension)
Subjective Ending Conditions	
SE1	Concise: the taxonomy is meaningful without being overwhelming
SE2	Robust: the dimensions and characteristics suffice to differentiate objects
SE3	Comprehensive: all objects can be classified
SE4	Extendible: new dimensions and characteristics can be added
SE5	Explanatory: the dimensions and characteristics explain an object

Our taxonomy and archetypes help answer the main research question: *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?* Because Nickerson et al. (2013) do not explicitly state how to induce or deduce concepts for a taxonomy, derive archetypes and evaluate the results, we extend their approach with other research methods. Four sub-questions guide us in the process of developing a taxonomy and deriving business model archetypes. In the subsequent paragraphs, the sub-questions and applied methods are briefly explained.

Sub-question 1: Based on an exploration of the business models of data marketplace owners in the B2B automotive industry and based on the existing generic business model taxonomies for data marketplaces, what dimensions can be derived to include in our preliminary business model taxonomy?

Business model literature about data marketplaces is scarce and comprehends predominantly theoretical concepts. As such, we take an explorative start to create our taxonomy. This is advised by Nickerson et al. (2013) when little data about the research domain is available. We follow the grounded theory method by Charmaz (2006), which is a suitable method for an explorative start. Charmaz (2006) proposes to construct theory through inductive reasoning, starting with information gathered from interviews, reports and other data materials instead of starting off with an academic literature review. Therefore, we start by conducting interviews with data marketplace owners in the B2B automotive industry. As explained in section 1.3, the B2B automotive industry is our industry of focus because it has established data marketplaces. By gathering data from interviews with data marketplace owners in this industry, we learn about the business models they apply in practice. Based on this information, we induce our first set of business model dimensions. Subsequently, these dimensions are supplemented with deduced dimensions from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) to create our preliminary taxonomy. Our preliminary taxonomy is still conceptual, which means that data marketplaces from the B2B automotive industry are not yet classified in the taxonomy.

Sub-question 2: What business model characteristics do owners of current data marketplaces in the B2B automotive industry apply?

To answer this question, we study the business models of a selection of data marketplace cases in the B2B automotive industry. Case documents are retrieved from online desk research. Mayring (2004) clearly describes a content-analysis method to induce concepts. We apply this method to induce business model characteristics from the case documents. Based on the induced characteristics, we refine our preliminary taxonomy. By classifying a selection of data marketplaces in the B2B automotive industry in our refined taxonomy, their business models can be distinguished. This is our application of the empirical-to-conceptual approach in Figure 4. We iterate this step until all objective ending conditions in Table 1 are met. The outcome of these iterations is our refined taxonomy.

Sub-question 3: What business model archetypes can be identified for data marketplaces?

At this stage, our taxonomy is designed. We use it to recognize patterns in the business models of the classified data marketplaces. Nickerson et al. (2013) do not explicate this step in their approach. By performing a cross-case analysis we recognize patterns in our taxonomy and create our business model archetypes. This method is commonly applied to search for patterns (Eisenhardt, 1989). We cluster data marketplaces with similar business model patterns in one archetype, following the pairwise comparison tactic by Eisenhardt (1989). Our archetypes can be used by researchers to analyze and develop business models of data marketplaces and by data marketplace owners to make design choices for their own business model.

Sub-question 4: Is our business model taxonomy evaluated as complete and useful?

Whether a taxonomy is complete and useful needs to be evaluated. According to Nickerson et al. (2013), a taxonomy is complete and useful when it satisfies all ending conditions from Table 1. During our taxonomy development process, we assess the objective ending conditions iteratively to decide when our taxonomy is finished. Next, the subjective ending conditions are assessed with experts during semi-structured interviews. Szopinski et al. (2019) identify expert interviews as a suitable method to evaluate whether a taxonomy is complete and useful. We conduct semi-structured interviews with Spiekermann (2019) and Fruhwirth et al. (2020) who we consider experts in the development of business model taxonomies for data marketplaces. Based on our assessment of the objective ending conditions and the feedback from the experts, we improve our taxonomy design.

An overview of our research process is provided in Table 2. Four main steps constitute the process in which we create a preliminary taxonomy, refine the taxonomy, apply our taxonomy and evaluate the results. A detailed description of our research process and applied methods are provided in chapter 3.

Table 2: Research process

Research step	Sub-question	Method	Data sources	Deliverable
Taxonomy conceptualization	*1: Based on an exploration of the business models of data marketplace owners in the B2B automotive industry and based on the existing generic business model taxonomies for data marketplaces, what dimensions can be derived to include in our preliminary business model taxonomy?*	Grounded Theory	Interviews and literature sources	Preliminary taxonomy
Taxonomy refinement	*2: What business model characteristics do owners of current data marketplaces in the B2B automotive industry apply?*	Desk research Content analysis	Case websites, whitepapers and external sources	Refined taxonomy

Pattern recognition	*3: What business model archetypes can be identified for data marketplaces?*	Cross-case analysis	Case websites, whitepapers and external sources	Business model archetypes
Evaluation	*4: Is our business model taxonomy evaluated as complete and useful?*	Semi-structured interviews	Expert interviews	Evaluated taxonomy and archetypes

1.6 CoSEM Relevance

Creating a business model taxonomy for data marketplaces concerns a design question for a complex socio-technical system. The social complexity lies in the number of parties involved in the trade of data and the organization of all trading participants at a data marketplace. The technical complexity lies in the features of data marketplaces such as the technical architecture and pricing mechanisms. Moreover, data trade requires institutional guidelines to regulate the trade of data and preserve data quality and privacy. The taxonomy and archetypes that are developed aid researchers and practitioners in Thus, technical, institutional, economic and social knowledge is required to gain understanding of business models of data marketplaces. Therefore, the topic meets the research objectives of the study program Complex Systems Engineering and Management.

1.7 Reading guide

The remainder of this thesis is structured as follows. The theoretical background of business models and trading structures is introduced in chapter 2. In this chapter we establish a theoretical foundation of concepts that return throughout this research. Next, the complete research method is described in chapter 3. In chapter 4, we present relevant business model categories for data marketplaces and derive business model dimensions to include in our preliminary business model taxonomy. Subsequently, the business model dimensions and characteristics based on six data marketplaces are described in chapter 5. These dimensions and characteristics compose our refined taxonomy. In chapter 6, we compare cases based on similarities and differences to generate business model archetypes. Next, the evaluation results of our taxonomy and business model archetypes are presented in chapter 7. In chapter 8, we discuss our research results. Finally, we conclude our research in chapter 9.

2. Theoretical Background

In this chapter, we provide theoretical background information about relevant concepts which return throughout this thesis. First, we introduce business model literature in section 2.1. Next, we discuss economic literature about different trading structures in section 2.2. Finally, we conclude the chapter in section 2.3.

2.1 Business Models

In this section the state-of-the-art in business models is introduced. Business models as defined by scientists are explained in chapter 2.1.1 to create consensus on how to represent the business model of data marketplaces. In chapter 2.1.2, we discuss business model components as defined in literature to generate main business model components that should be included in the formation of our taxonomy.

2.1.1 Business Model Definition

The missing consensus on a universally accepted business model definition creates uncertainty about the representation of a business model. Chesbrough & Rosenbloom (2002) describe business models as frameworks that convert technological input into economic output. Hence, business owner can use business models to transform technical potential into economic value. Teece (2010) defines business models as the design of value delivery to the customer. Amit & Zott (2001) explain that business owners use business models to visualize the design of “transaction content, structure and governance”, create value from various sources and discover new business opportunities. They continue stating that the lack of geographical borders and high volume of information goods in virtual markets, cause traditional industries to change. Despite the variety in business model definitions, all business model descriptions include component-based perspectives (Hartmann et al., 2014). Therefore, we represent business models based on the business model components in this study. In section 2.1.2 we aim to identify the business model components in the most cited business model literature and select the main business model components that form the basis of our business model taxonomy.

2.1.2 Business Model Components

Table 3 shows some of the most cited authors in the business model literature. They identify business model components that are applicable to any organization. In addition to these highly cited papers, we consider business model components identified in the STOF model by Bouwman et al. (2008). Bouwman et al. (2008) focus on business models for ICT-enabled services in their STOF model, which contains technical aspects. Data marketplaces have a technical infrastructure and deliver ICT-enabled services. Therefore, the technical aspects are relevant and need to be represented in the business model. Table 3 provides an overview of the components that we identified.

Table 3: Business model literature

Year	Author	Cited	Business model components
2010	Chesbrough	3686	Value proposition, market segment, value chain, cost structure and profit potential, revenue mechanisms, value network, competitive strategy
2010	Teece	7046	Select technologies and features, benefit to customer, market segments, revenue stream, capture value
2010	Osterwalder & Pigneur	11182	Customer segment, value proposition, channels, customer relationship, revenue streams, key resources, key activities, key partners, cost structure
2002	Chesbrough & Rosenbloom	5482	Value proposition, market segment, value chain, cost structure and profit potential, value network, competitive strategy
2001	Amit & Zott	7175	Novelty, lock-in, complementary, efficiency
2008	Bouwman et al.	449	Services, technology, organizational arrangements, finance

We classify the business model components under the main components; value creation, value delivery and value capture (Teece, 2010). These are the meta-characteristics in our business model taxonomy. First, value creation is the process of making something that brings worth to the customer. The components value proposition, customer segment and customer relationships are assigned to this meta-characteristic. The **value proposition** comprises the product or service offering (Chesbrough, 2010; Osterwalder & Pigneur, 2010; Teece, 2010). A company creates a value proposition to solve a customer problem. The value proposition is designed for a group of customers from a market segment (Chesbrough & Rosenbloom, 2002). Osterwalder & Pigneur (2010) define this as the **customer segment**. Additionally, they state that companies influence the overall customer experience by creating **customer relationships**. These relationships range from personal assistance with a high level of human interaction to automated services that are performed online with minimal human interaction. With the formation of customer relationships, companies aim to acquire customers and increase sales.

Second, value delivery is about the asset arriving at the customer. Chesbrough & Rosenbloom (2002) conceptualize the value chain that distributes the value offering. The value chain comprises the processes, activities, relevant resources and capabilities required to build and distribute the proposition. Bouwman et al. (2008) and Chesbrough & Rosenbloom (2002) mention the value network that defines relationships with other partners considering the supply of resources. Overall, four main components can be identified in the value chain and value network, acknowledged by Osterwalder & Pigneur (2010). These are the channels, key resources, key activities and key partners which we assign to the meta-characteristic value delivery. Companies communicate, distribute and sell their value proposition through their **channels**. The channels are the customer-company interface through which customers purchase the products or services. Companies produce and deliver their value proposition using **key resources**. These be physical, financial, intellectual and human resources. Examples of the **key activities** that companies perform are the production of the value proposition, maintenance of the channels and training of employees. A firm must perform these actions to operate their business model. A firm relies on **key partners** to provide their service. Partnerships are formed to outsource certain operations. For example, a firm can choose to outsource human resource management to a company who specializes on this front. By outsourcing operations that are not within their area of expertise, firms can reduce costs.

Third, when companies capture value, they monetize the created and delivered value. Many businesses assume that when they create a product or service, the customer will pay for it (Teece, 2010). According to Teece (2010), this is a common mistake made in markets. Companies sell their technological invention, instead of a solution that the customer needs. It is unlikely that companies can capture value from items that do not solve a problem. Capturing value from the trade of intangible goods is also problematic. Property rights of intangibles are unclear, which makes its pricing difficult (Teece, 2010). The trade of such goods requires a revenue model that captures value from the sale a solution, not an item. The revenue model includes the revenue streams and the pricing model. Osterwalder & Pigneur (2010) distinguish **revenue streams**, such as licensing and brokerage, and **pricing models**, like fixed and dynamic pricing. These components are assigned to the meta-characteristic value capture. Value capture also includes the **cost model** that covers all expenses of a company to operate the business model.

The business model canvas by Osterwalder & Pigneur (2010) serves as leading framework to identify the main business model components. The business model canvas combines all aspects that are identified by Amit & Zott (2001), Chesbrough & Rosenbloom (2002) and Teece (2010). We also consider characteristics from the STOF model by Bouwman et al. (2008). They define data as a resource to deliver a value proposition. *Data* is the intermediate good that is exchanged at a data marketplace and is key in their value proposition. According to Bouwman et al. (2008), companies build the architecture of a digital service with technical resources. Because data marketplaces perform digital services, we include the resource characteristic *technical*. Table 4 shows an overview of the main business model components and their characteristics from Osterwalder & Pigneur (2010) to which we add the characteristics data and technical from Bouwman et al. (2008). The components will be further refined throughout this research and business model characteristics will be specified for data marketplaces. Some characteristics may not be applicable to data marketplaces and are not included in the final taxonomy.

Table 4: Business model components adapted from Bouwman et al. (2008) and Osterwalder & Pigneur (2010)

	Component	Description	Characteristic	Description
Value creation	Customer segment	The different groups of people or organizations served by an enterprise	Mass market	There is no distinction between customer segments
			Niche market	Specialized products or services for specific customer segments
			Segmented	Markets are segmented to customers with slightly different needs
			Diversified	Unrelated customer segments are served
			Multi-sided platforms	Interdependent customer segments are served
	Value proposition	The bundle of products and services that create value for a specific Customer Segment	Data	Data is characterized by its transfer time and volume
			Newness	A complete new set of needs is served
			Performance	Product or service performance is improved
			Customization	Products or services are tailored for specific customers
			Getting the job done	Customers rely on the product or service to do something
			Design	The design of a product or service makes it stand out
			Brand/status	The brand gives the customer a certain status
			Price	Competing at a lower price satisfies price-sensitive customers
			Cost reduction	Helping customers reduce costs
			Risk reduction	Customers experience lower risk by purchasing the product or service
			Accessibility	Access to products or services is granted which it was not available before
			Convenience	Things are made more convenient or easy to use
	Customer relationship	The types of relationships a company establishes with specific Customer Segments	Personal assistance	Customer assistance happens through human interaction
			Dedicated personal assistance	Individual clients are helped by a dedicated customer representative
			Self-service	With the means of the company, the customers can help themselves
			Automated services	Automated services recognize customers to offer customized services
			Communities	Customers are connected via communities
			Co-creation	Value is co-created between the business owner and its customers
Value delivery	Channels	How a company communicates with and reaches its Customer Segments to deliver a Value Proposition	Own	Channels are owned by the business owner
			Partner	Channels are owned by a partner
			Direct	Channels are directly operated by the business owner
			Indirect	Channels are indirect and offer a range of options
	Key resources	The most important assets required to make a business model work	Physical	Physical, often capital-intensive resources
			Intellectual	Proprietary knowledge
			Human	People as the main resource
			Financial	Financial guarantees
			Technical	Applications, devices, service platforms, access networks and the backbone infrastructure form the technical architecture of services
	Key activities	The most important things a company must	Production	Producing substantial quantities
			Problem solving	Knowledge management or continuous training

		do to make its business model work	Platform/network	Maintaining the platform; e.g. platform management, service provisioning and platform promotion
	Key partners	The network of suppliers and partners that make the business model work	Strategic	Alliances between non-competitors
			Coopetition	Alliances between competitors
			Joint venture	Alliances to develop new businesses
			Buyer-supplier	Alliances to assure supplies
Value capture	Revenue streams	The money a company generates from each Customer Segment	Asset sale	Ownership rights are sold to a physical product
			Usage fee	The customer pays for the use of a service
			Subscription fees	Continuous access to a service is sold
			Lending	Access to a service or product is granted temporarily
			Licensing	Permission is granted to use intellectual property
			Brokerage fee	Intermediaries act on behalf of two or more parties
			Advertisement	Revenue is generated by letting other brands advertise
	Pricing model	The pricing mechanisms of the sold product or service	Fixed	List price, product feature/customer segment/volume dependent
			Dynamic	Negotiation, yield management, real-time-market, auctions
	Cost model	All costs incurred to operate a business model	Fixed costs	Independent of the volume, the costs stay the same
			Variable costs	Costs vary with the volume
			Economies of scale	When the outputs increase, costs decrease
			Economies of scope	Larger operations cause lower costs

2.2 Trading Structures

Data marketplaces differ in their orientation. As explained in section 1.2.1, we define data marketplace types with a hierarchical and market orientation. In this section fundamental economic theories are discussed to characterize the orientation structures. Williamson (1973; 1989) established the market – hierarchy continuum to explain factors that cause organizations to shift from a market to a hierarchical structure. We use those factors to characterize the hierarchical and market orientation of data marketplaces in section 2.2.1. Next, the network theory developed by Powell (1990) is introduced in section 2.2.2. He criticizes the market – hierarchy continuum and argues that there are organizations that are neither market nor hierarchically structured. We introduce his theory to clarify that our focus on the hierarchical and market orientation may not provide a satisfactory overview of economic exchange. The theory of Powell (1990) on network structures enables us to reflect on our definition of data marketplaces and their orientation structures in discussing the results. This is further explained in section 2.2.3.

2.2.1 Market – Hierarchy Continuum

Williamson (1973) presents the hierarchical and market structures as opposites. According to Williamson (1973) there are factors that cause transactions to shift from markets into hierarchies. These factors contribute to market failure. He discusses human factors and transactional factors that lead to unfavorable market conditions. The human factors are (i) bounded rationality and (ii) opportunism. The transactional factors are (i) uncertainty and (ii) small numbers. Figure 5 shows the human and transactional (environmental) factors as visualized by Mahoney (2004). In the following paragraphs we further clarify these factors.

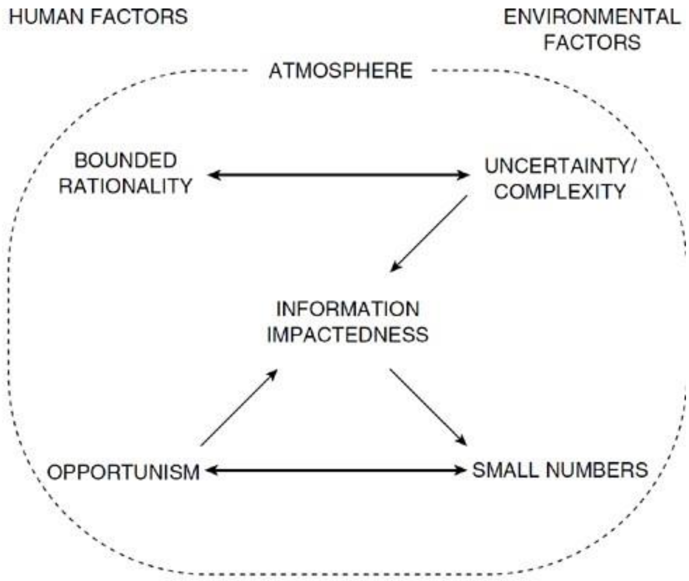

Figure 5: Human and transactional factors (Mahoney, 2004, p. 122)

First, bounded rationality is the inability of humans to "receive, store, retrieve and process information without error" (Williamson, 1973, p. 107). Although humans try to act rational, with limited information they reach a satisfactory solution instead of an optimal solution. In a hierarchical structure, bounded rationality poses less of a problem than in a market structure. The clear boundaries in departments, lines of authority and reporting mechanisms internalize transactions in a hierarchical structure. This enables these organizations to "write contracts that cover all possible contingencies" (Powell, 1990, p. 297). In a market structure, boundaries are less clear. Bounded rationality in organizations with a market structure makes organizations prone to market failure.

Second, opportunism concerns the aim of actors to maximize their personal gain (Williamson, 1973). People can go as far as deceit to achieve their goal. Powell (1990) explains that actors aim to minimize their costs in organizations with a market structure. Production and exchange at these organizations are determined by price competition. When the price of a product does not satisfy the needs of an actor in a market structure, he has the flexibility to move to another seller who does meet his requirements. Powell (1990) notes that interactions in a market structure do not "establish strong bonds of altruistic attachments" (p. 302). This leads to quick and efficient interactions with a lack of strong relationships. On the other hand, in hierarchical structures actors practice authority by imposing rules and sanctions to regulate opportunistic behavior (Williamson, 1989). According to Powell (1990), actors communicate based on routines with people they are familiar with. As such, people who "know one another, have a history of previous interactions and possess a good deal of firm-specific knowledge" trade in hierarchical structures (Powell, 1990, p. 302). Authoritative relations and personal identification pose less room for opportunistic behavior in a hierarchical structure. According to Williamson (1973), opportunism does not lead to the success of one structure over the other. Opportunism indicates what actors are more likely to trade in what structure. People who aim for maximum personal gain are attracted to a market structure and people who seek routine are attracted to a hierarchical structure.

Third, uncertainty in a market structure influences economic behavior (Williamson, 1973). Examples of uncertain factors at the time of exchange in a market structure are future price, demand/supply ratio and price/quality estimation. This may lead to non-optimal transactions. Powell (1990) explains that there is more control over the coordination supply and demand in a hierarchical structure, stating "the visible hand of management supplants the invisible hand of the market in coordinating supply and demand" (p. 303). He explains that managerial teams in a hierarchical structure have the ability to coordinate high volume and speed operations. The vertical integration of organizations with a hierarhical structure enables them to well-coordinate mass production and distribution. This mitigates uncertainties.

Fourth, small numbers are unfavorable for a market structure (Williamson, 1973). A high number of buyers and sellers creates a competitive environment in market structures, which stimulates price reduction. Small numbers cause an organization with a market structure to shrink or vanish altogether.

In addition to the human and transactional factors, Williamson (1989) introduces asset specificity. Asset specificity entails the extent to which an asset can be used for multiple purposes. Williamson (1989) recognizes five different forms of asset specificity. These are site specificity, physical asset specificity, human asset specificity, dedicated assets and brand name capital. Powell (1990) elaborates that "transaction-specific investments – of money, time and energy that cannot be easily transferred – are more likely to take place within hierarchically organized firms" (p. 297). Thus, assets that are more knowledge specific are likely traded in organizations with a hierarchical structure.

We use the previous notions of asset specificity and opportunism to define organizations with a hierarchical or market structure. Asset specificity is a relevant concept for the value proposition of a data marketplace. *Asset specificity* concerns the goods that are traded in a hierarchical or market structure. Organizations with a hierarchical structure likely trade in asset specific. Organizations with a market structure trade in less asset specific goods. Opportunism is a relevant factor to describe the customers who are attracted to a data marketplace. *Opportunism* concerns the people who are attracted to organizations with a hierarchical or market structure. Organizations with a hierarchical structure trade on authoritative basis between actors who are familiar with each other. Organizations with a market structure trade on competitive basis between actors who aim for the maximum individual gain. This leads to our definitions of organizations with a hierarchical and market orientation:

i. *Organizations with a hierarchical orientation* trade in asset specific goods, on authoritative basis between actors who are familiar with each other
ii. *Organizations with a market orientation* trade in less asset-specific goods, on competitive basis between actors who aim for the maximum individual gain

2.2.2 Network Structure

Organizations do not always orient themselves in either a market or hierarchical structure. Powell (1990) criticizes the market-hierarchy continuum of Williamson (1973;1989) because it does not provide a complete perspective on trading structures. He disagrees with the sharp boundaries of the hierarchical structure on one hand and the market structure on the other hand. According to Powell (1990), there are organizations with less clear boundaries that transact on collaborative basis. These organizations do not resemble market structures or hierarchical structures. Therefore, he calls for the inclusion of a third structure: the network structure. Table 5 shows a comparison of the trading structures as defined by Powell (1990). In the following paragraphs we characterize the goods traded by and actors attracted to organizations with a network structure to define organizations with a network structure.

Table 5: Comparison of trading structures (Powell, 1990, p. 300)

	Forms		
Key features	*Market*	*Hierarchy*	*Network*
Normative basis	Contract – property rights	Employment relationship	Complementary strengths
Means of communication	Prices	Routines	Relational
Methods of conflict resolution	Haggling – resort to courts for enforcement	Administrative fiat – supervision	Norm of reciprocity
Degree of flexibility	High	Low	Medium
Amount of commitment among the parties	Low	Medium to high	Medium to high

Tone or climate	Precision and/or suspicion	Formal, bureaucratic	Open-ended, mutual benefits
Actor preferences or choices	Independent	Dependent	Interdependent

The goods traded in network structures are different from the goods traded in hierarchical or market structures. Powell (1990) finds that network structures are especially suitable to trade goods "whose value is not easily measured" (p.304). For example, know-how goods are difficult to trade in market structures and inefficiently communicated in hierarchical structures (Powell, 1990). Actors in network structures do not necessarily trade goods to make profit, but are motivated by the gain of new information and skills. Therefore, network structures are suitable to trade assets whose value is difficult to measure.

The actors who trade in network structures are driven by different motives than actors in hierarchical or market structures. Relationships, mutual interest and reputation are of interest to actors in network structures (Powell, 1990). The complementary strengths, that the trading parties in a network have, help them solve problems together. The actors have the common goal to maintain a good reputation and establish friendships. This creates interdependent relationships between the actors who need to collaborate to maintain their network. However, Powell (1990) notes that trade in a network structure is not solely based on collaboration: "each point of contact in a network can be a source of conflict as well as harmony" (p. 305). For example, companies are rivals in terms of market share in strategic alliances, but collaborate to increase economies of scale. The emergence of access restrictions is apparent in a network structure. The repeated trading between actors who are connected raises barriers of entry. Newcomers will experience a harder time to create new relationships and enter the network. Overall, actors in a network structure collaborate for mutual benefits and are strongly connected.

In short, certain circumstances call for an alternative trading structure next to hierarchies and markets. The assets that are traded in network structures differ from hierarchical and market structures. Organizations with a network structure trade in assets whose value is difficult to measure. The people who are attracted to a network structure differ from hierarchical and market structures as well. Organizations with a network structure trade on collaborative basis between actors who aim for to achieve mutual benefits. This results in our definition of a network trading structure:

Organizations with a network structure trade in assets whose value is difficult to measure, on collaborative basis between actors who aim to achieve mutual benefits.

2.2.3 Implications for Data Marketplace Types

In the previous sections we explained that organizations have various trading structures. In Williamson's (1973; 1989) view, organizations are oriented towards either a hierarchical structure or a market structure. This results in a continuum in which data marketplaces can be classified as hierarchically oriented or market oriented as visualized in Figure 6.

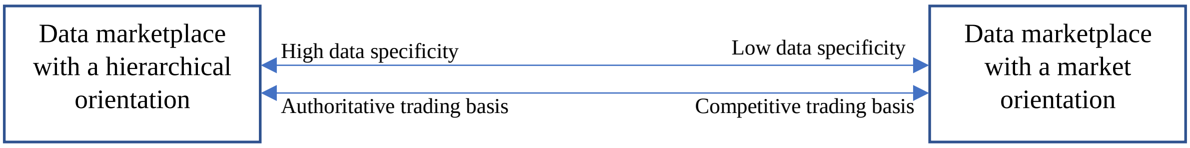

Figure 6: Data marketplaces in the market-hierarchy continuum

According to Powell (1990), this continuum is too narrow. He argues that many organizations have alternative trading structures and introduces the netwrk structure. When defining the orientation of data marketplace types according to the view of Powell (1990), three data marketplaces with distinctive orientations exist (see Table 6).

Table 6: data marketplaces as three distinctive trading structures

Data marketplace with a hierarchical orientation	Data marketplace with a market orientation	Data marketplace with a network orientation
• Data is asset specific • Authoritative, trading basis between actors who are familiar with each other	• Data is less asset specific • Competitive trading basis, between actors who aim for maximum individual gain	• Data value is difficult to measure • Collaborative trading basis between actors who aim to achieve mutual benefits

In our definition of data marketplaces we choose to continue with the market-hierarchy continuum to research data marketplace types. This decision is based on the presence of two distinctive orientations of data marketplaces in practice and literature. In practice, data marketplaces with a hierarchical orientation exist. These data marketplaces trade on bilateral basis (Koutroumpis et al., 2017). Attempts are made to launch multilateral data marketplaces with a market orientation in practice, but these initiatives do not succeed. Examples of such data marketplaces that withdrew from the market are Microsoft Azure Data Marketplace, Kasabi and InfoChimps (Spiekermann, 2019). In literature, scientists focus on data marketplaces with a market orientation to advance the development of these data marketplace types (Fruhwirth et al., 2020; Schomm et al., 2013; Spiekermann, 2019). In line with the data marketplace classifications of Stahl et al. (2016) and Koutroumpis et al. (2017), we research data marketplaces in the market – hierarchy continuum to distinguish between data marketplaces in practice and theory.

The network structure is not included as a third orientation in our definition of data marketplaces, because we observed a market – hierarchy continuum in practice and literature. However, the view of Powell (1990) is not completely discarded in this thesis. Powell (1990) makes the valid argument that some organizations are neither hierarchically nor market oriented. It is plausible that data marketplaces have a network orientation. First, participants at data marketplaces trade in data, a good that is difficult to price (Koutroumpis et al., 2017; Powell, 1990). Assets whose value is difficult to measure are commonly traded in a network structure. Second, data marketplaces enable data sellers, data buyers and third party service providers to trade data with each other. Thomas & Leiponen (2016) describe data exchange among enterprises as a form of collaboration. Collaboration is another characteristic of organizations with a network structure. In chapter 8.5, in the discussion, we reflect whether the market-hierarchy continuum suffices to describe data marketplace types or whether there is a need to include the network orientation as a third trading structure in future research.

2.3 Conclusion of Chapter 2

Our objective is to clarify what business models owners of different types of data marketplaces in the B2B automotive industry apply. In chapter 2.1 we identified business model components for our business model taxonomy. The customer segment, value proposition, customer relationships, channels, key resources, key activities, key partners, revenue, pricing model and cost model are part of a business model and need to be specified for data marketplaces in our taxonomy. The data marketplace types that are considered in this research differ in hierarchical or market orientation. This is in line with the data marketplace classifications of Stahl et al. (2016) and Koutroumpis et al. (2017) and the market–hierarchy continuum as defined by Williamson (1973; 1989). However, the market–hierarchy continuum may not provide a complete view on data marketplace types. According to Powell (1990), many organizations have a network structure, distinctive from a hierarchy or market structure. In chapter 8.5 we reflect whether the market-hierarchy continuum is satisfactory to describe data marketplace types or whether future research needs to consider the network orientation.

3. Research Methods

By means of qualitative research, we aim to identify business models of different types of data marketplaces. This is achieved by the creation of a taxonomy, the classification of data marketplace cases within this taxonomy and finally the evaluation of the taxonomy. Taxonomies are designed early on in a research process to contribute to theory building (Bapna et al., 2004; Nickerson et al., 2013; Szopinski et al., 2019). The development of taxonomies is essential to advance business model literature for data marketplaces. The domain that we research is the business model of data marketplaces in the B2B automotive industry.

In chapter 1.5, the taxonomy development method of Nickerson et al. (2013) is introduced. Because they do not specify methods how to induce and deduce components, we extended their iterative inductive and deductive approach with other research methods. We started our taxonomy development process with the application of the Grounded Theory method to induce taxonomy dimensions from interviews with practitioners. These dimensions were aligned with dimensions that are deduced from the business model taxonomies developed thus far. Next, the characteristics within the taxonomy were induced by applying content analysis. Followed by cross-case analysis to identify the business model archetypes. Finally, through conducting semi-structured interviews the taxonomy was evaluated. An overview of the objectives, methods and outputs is provided in Figure 7. In the subsequent paragraphs, we discuss these methodologies more specifically.

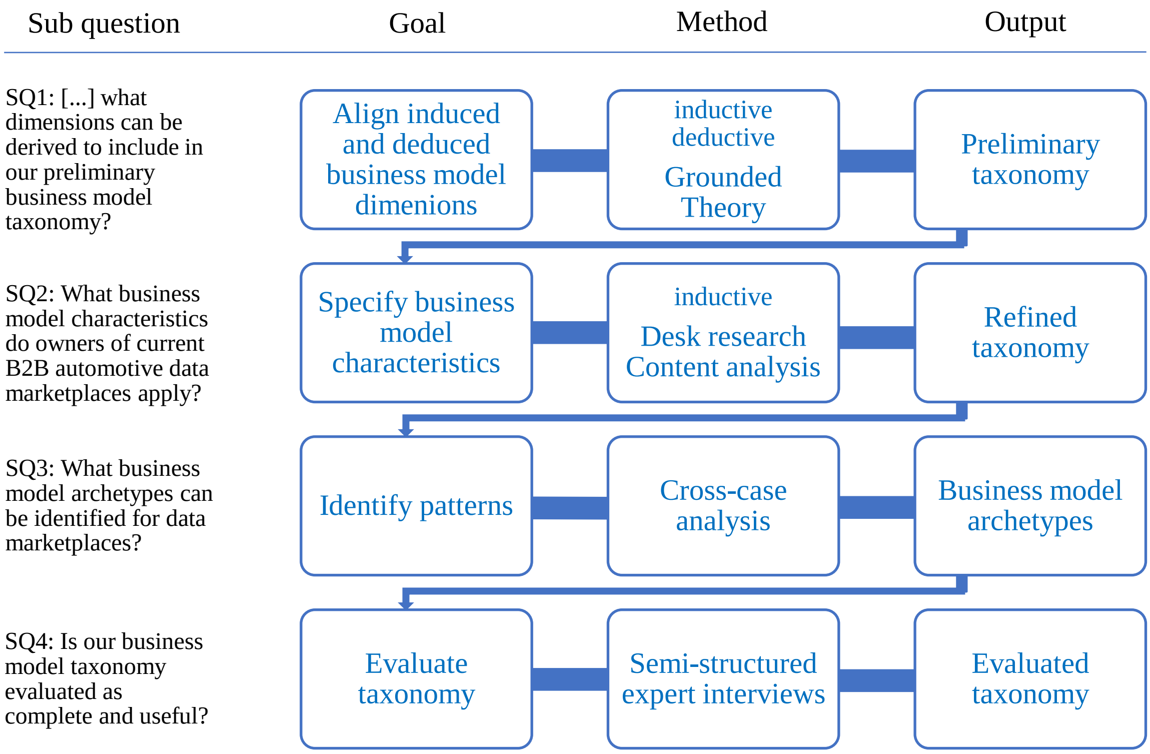

Figure 7: Overview research methods

3.1 Grounded Theory

In this section, our application of the Grounded Theory method is explained. The goal in this step is to induce and deduce business model dimensions to answer the first sub-question: *Based on an exploration of the business models of data marketplace owners in the B2B automotive industry and based on the existing generic business model taxonomies for data marketplaces, what dimensions can be derived to include in our preliminary business model taxonomy?* Our preliminary taxonomy is based on induced dimensions generated with the application of the Grounded Theory method and deduced dimensions from the existing taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020).

Generating the required business model insights for data marketplaces is challenging, because business model literature about data marketplaces is scarce. The limited available literature demands an explorative start in which we gather data from other sources. Therefore, we followed the Grounded theory method by Charmaz (2006). According to Charmaz (2006), grounded theory is constructed through inductive reasoning, starting with information gathered from interviews, reports and other data materials instead of starting off with an academic literature review. She suggests to develop fresh theory by starting the analysis with interview data. Hence, we started our taxonomy development process by conducting interviews with data marketplace owners to learn about their business models. In the second step of constructing grounded theory, we coded the interview data and constructed categories. Coding interview data early in the data collection process forces researchers to directly start analyzing the data and recognize gaps in emerging theory (Charmaz, 2006). Based on new discoveries during the coding practice, we revisited our interview data and conducted more interviews to fill gaps in our data. In the third step, the business model categories were enriched with technical literature (e.g. research studies and theoretical papers) and nontechnical literature (e.g. manuscripts, records, reports). Strengthening our categories with extant literature helps to support our analytical arguments (Charmaz, 2006). Based on new discoveries in the literature sources, we revisited our interview data and categories to explore new ideas. After the enriched categories were created, we induced business model dimensions to include in our preliminary taxonomy. Last, the induced dimensions were aligned with deduced dimensions from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) to create our preliminary taxonomy. We did not start with deducing dimensions, because two existing taxonomies of data marketplaces do not provide a sufficient amount of data to base our taxonomy on. By gathering data from interviews first, we learned about relevant business model dimensions for data marketplace owners in practice and supplemented these with dimensions from theory.

Overall, four main steps were performed to iteratively construct our preliminary taxonomy. These are (i) conduct interviews, (ii) construct categories, (iii) enrich categories and (iv) align dimensions, visualized in Figure 8. In the following sections we explain these steps in more detail.

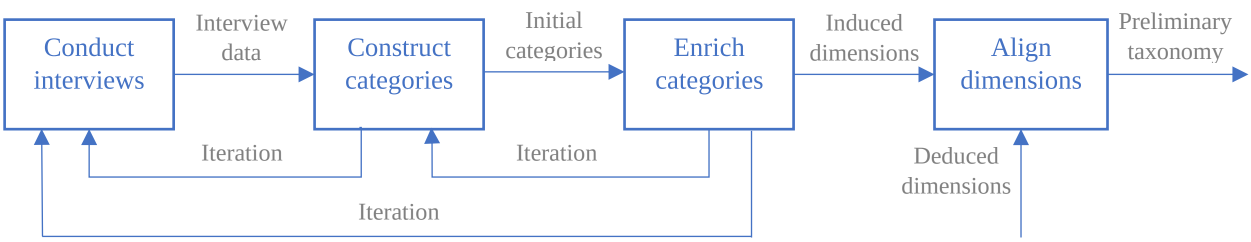

Figure 8: Process Grounded Theory

3.1.1 Conducting Interviews

By conducting interviews with data marketplace owners in the B2B automotive industry, we tried to identify business model components that they apply in practice. In grounded theory, the involvement of academic literature is deliberately delayed to avoid the formation of theory through extant ideas (Charmaz, 2006). In our selection of interviewees, we aimed for variance in the data marketplace types that the interviewees represent (see Table 7). The types range from hierarchical to market orientation and private to independent ownership, as introduced in section 1.2.1. At minimum, each marketplace type has to be represented by at least one interviewee. In addition, we maintained the following selection criteria:

- The interviewees are available for an interview and speak English
- The interviewees work at data marketplaces that trade in automotive data.
- The interviewees have business model knowledge. We judged this based on their job title. Interviewees who occupy a position related to business development, are expected to have in-depth knowledge about the business model.

- The interviewees have over 5 years of work experience in business development or consultancy. Because many data marketplaces are newly founded in the last 5 years, we could not expect the interviewees to have over 5 years work experience at the respective data marketplace. Therefore, we looked at their work experience previous to their current job.

Table 7: Overview of respondents for the intensive interviews

Code	Type	Job title	Other relevant experience
DM1	Market, consortium	Business development	Previously worked as a marketing and business development consultant for 4 years
DM2	Market, independent	Product owner	Over 5 years of experience as a data scientist and business consultant at various multinationals
DM3	Market, independent	Unknown	7 years of experience in advising ministries about traffic and mobility data
DM4	Market, independent	Business Development	Over 5 years of experience as a freelance consultant
DM5	Market, government	Innovation Manager Smart Mobility	Previously worked as a consultant for national agencies and has over 5 years of experience working on smart mobility projects
DM6	Market, private	Director Business Development	Over 10 years of experience at various IT service providers as sales manager
DM7	Hierarchical, private	Head of Enterprise Business Development	Over 8 years of experience in corporate development at a multinational

Intensive interviews were conducted to explore new perspectives on business models for data marketplaces. During intensive interviewing we encouraged, listened to and learned from the interviewees. Charmaz (2006) suggests this interview method to construct grounded theory. She argues that intensive interviewing is a suitable technique to explore a topic in-depth and gather rich data. Therefore, a study of seven interviews that allow for in-depth exploration of a topic adds more value than twenty interviews that do not cover the underlying problems. The goal of the interviews was to facilitate open ended and in-depth exploration of the interviewees' area of expertise and construct conceptual categories. The categories represent components that need to be considered for the business model of a data marketplace. Charmaz (2006) developed a theory building process that starts with open ended interview questions. She suggests a number of initial and ending guiding questions to gather rich data (Charmaz, 2006, p. 30-31). Based on those suggestions, we formulated four initial questions to start the conversation:

1. Could you describe the main trends for data marketplaces and how does [name data marketplace] respond to those trends?
2. Could you describe the main challenges data marketplaces face and how does [name data marketplace] respond to those challenges?
3. Could you explain the key components of the [name data marketplace] business model?
4. Could you describe the difference between [name data marketplace] and [name competitor]?

Based on the responses to the initial questions, intermediate questions were formulated as the interviews proceeded. This results in a diverse range of follow-up questions, which caused no interview to be the same. For example, one interviewee described data privacy as a main challenge for data marketplaces. Data privacy forms an issue because people are reluctant to share their data with a company. People are willing to share their own data on social media, but are uncomfortable with directly sharing their data for commercial purpose. Subsequently, we asked the follow-up question; "how does [name data marketplace] incorporate privacy preservation in their business model?" By asking 'how' questions, we tried to elicit rich data. The interviewee explained that they see privacy regulation as a business opportunity. Privacy regulation is something their customers are less familiar with. Their data marketplace functions as a consent management hub to take the privacy concern away from their customers. Data is only traded with consent from the data owner. Another interviewee responded that the pricing of data forms a challenge for data marketplace owners and their customers. Consequently, the interviewee was asked the following question "How does the data marketplace support their participants in capturing value from data?" The interviewee answered that they educate their customers about possible pricing strategies. In online documentation, the data marketplace owner provides information about data pricing strategies. The data sellers have the final say in pricing their data. The interviewee acknowledged that there is no dynamic pricing mechanism to properly capture value from data yet. The initial question on challenges led to two different answers. By carefully listening to the responses of the interviewees and asking follow-up questions we were able to gather diverse data about business models of data marketplaces.

3.1.2 Constructing Categories

Constructing categories is the first step in interpreting the interview data. We coded the data qualitatively to construct the data into categories that represent business model concepts. Charmaz (2006) describes qualitative coding as "the process of defining what the data are about" (p. 43). During this step the process of data collection is linked to theory building. Charmaz (2006) identifies two main coding phases that we followed: initial coding and focused coding.

First of all, initial coding involved naming all data segments. Charmaz (2006) introduces three different approaches for initial coding. These are word-by-word coding, line-by-line coding and incident-to-incident coding. Word-by-word analysis is a nuanced form of coding and is rather applied to internet data. This approach would be suitable for short survey answers. Line-by-line analysis is generally applied to data about empirical problems or processes such as observations and autobiographies. For incident-to-incident analysis, codes are compared to incidents coded earlier. This approach suits behavioral observations with little additional context. We applied line-by-line coding for initial coding because the interviews that we conducted are too extensive for word-by-word coding. Incident-to-incident coding was also not applied, because our interview data does not cover behavioral observations, but business processes. The description of fundamental processes makes the interview data suitable for line-by-line coding.

For example, one interviewee described their data processing activities saying *"Data marketplaces often need to do data aggregation before giving data to the user. We use data cataloguing for this process. There you can do data tagging and data cleansing"* (DM2). We assigned the following codes to these lines: *aggregate data*, *catalogue data*, *tag data* and *clean data*.

Second, during focused coding the most frequent or significant codes were selected to create categories that cover larger segments of text. Focused coding is the process during which the researcher starts to recognize relationships and patterns between categories (Charmaz, 2006). First, we created focused codes per interview transcript. For example, based on the initial codes from of the interviewee describing the data processing activities (as described above), we selected *aggregate data* as the most significant code for this piece of text. After we completed this process for all interviews separately, a second round of focused coding was performed to construct categories that apply to all interviews. This second step of focused coding was required, because interviewees use different wordings to describe similar processes. For example, the focused codes *searching databases*, *aggregate data* and *harmonize and synchronize data* all refer to the data processing activities that a data marketplace owner performs. Hence, in the second round of focused coding, we created the overarching category *data processing activities*. Likewise, six more categories were formed; *data regulation*, *customers*, *platform infrastructure*, *revenue model*, *data quality* and *other* (see Table 8).

Table 8: Constructed categories

Main categories	Focused codes
Data regulation	Design quality standards / smart contract / comply with GDPR / delegated data regulation / preserve data privacy / comply with EU law / setting legal framework is challenging / terms and conditions / privacy is a challenge / privacy regulation / privacy challenge / check privacy regulation / terms and conditions determine data usage / Privacy disables open data publication / use of data is a license / non-cooperation of OEMs / customers restrict data usage / data ownership
Customers	Users / industry domain / attracting a specific customer segment / large target group / customer segment / maintain customer segments / direct OEM relationship
Platform infrastructure	Decentral data control / open governance / centralized or decentralized approach / open protocol / open platform infrastructure / decentral infrastructure / decision making at consumer / challenge to regulate IT integration
Data processing activities	Searching databases / saving time / overview in catalog / perform additional activities / provide corporate and open data / aggregate data / perform activities for all needs / advertise meta data / mixed functionalities / preserve data privacy / national access point / offer broker services / regulate data availability / broker of data / harmonize and synchronize data / setting legal framework is challenging / data processing activities / acting as traditional marketplace / extracting value from data is a challenge / acting as a consent management hub / key activities / key processes / perform several activities / advise OEM in data supply / market leader for development and research / develop own data products / perform activities / enable analysis of car data / generating insights from sensor data is difficult / collect data which is needed / differentiate added value / expand product offering / value chain depends on layer / multiple suppliers cause more activities / production is partially standardized and partially customized / try to standardize terms and conditions / high service quality / added value of aggregated product / performance of data processing activities
Revenue model	Valorizing data is difficult / no active role in data pricing / price discovery / explaining the crypto currency is a challenge / pricing / online product prices / data licensing / Licensing disables open data publication
Data quality	No check of data quality / cooperate to create quality standards / cooperate to automatically improve data
Other	Data marketplace concept / small company size / explain data marketplace concept / role MDM / create fit between customer and marketplace / data marketplace type / fit between governance and client / increase in data generation / more hardware in vehicles / evolve mobility definition / cooperate to develop solution / More data collection because of partnerships

3.1.3 Enriching Categories

The categories *data regulation*, *platform infrastructure*, *data processing activities* and the *revenue model* were main topics of discussion during the interviews. We further enriched these categories with technical and non-technical literature during theoretical sampling. Theoretical sampling is the process of collecting data from technical and non-technical literature to develop our tentative categories from the previous step into theoretical categories (Charmaz, 2006; Strauss & Corbin, 1998). We were able to fill gaps and strengthen our categories with concepts from extant literature. Table 9 shows the literature sources that we used for theoretical sampling. The technical literature is shown with code TL. The nontechnical literature sources, represented with code NTL, are reports written by consultancies and governmental research organs about automotive data marketplaces. We refer to these sources in chapter 4 when we discuss our enhanced categories.

Table 9: Technical and nontechnical literature

Code	Author(s)	Used to enhance category…
TL1	Allee (2008)	Data processing activities
TL2	Christiaanse & Markus (2002)	Customer relationships
TL3	Curry (2016)	Data processing activities
TL4	Janssen & Verbraeck (2005)	Data processing activities
TL5	Koutroumpis et al. (2017)	Data regulation, customer relationships, platform infrastructure, data processing activities
TL6	Koutroumpis et al. (2020)	Platform infrastructure
TL7	Muschalle et al. (2012)	Revenue models
TL8	Nicolaou & McKnight (2006)	Customer relationships

TL9	Ølnes et al. (2017)	Data regulation, data processing activities
TL10	Pavlou (2002)	Customer relationships
TL11	Savelyev (2017)	Data regulation
TL12	Schomm et al. (2013)	Revenue models
TL13	Thomas & Leiponen (2016)	Data processing activities
TL14	Truong et al. (2012)	Data regulation
NTL1	Bertoncello et al. (2016)	Customer relationships
NTL2	Deichmann et al. (2016)	Data processing activities
NTL3	Duch-Brown et al. (2017)	Data regulation, revenue models
NTL4	Martens & Mueller-langer (2018)	Customer relationships, data processing activities, revenue model
NTL5	Ramirez et al. (2014)	Data regulation

An example of a dimension that is induced based on our enriched categories is the dimension *contract*. Initially, we constructed the tentative category data regulation. The interviewees explained that they need to comply to privacy regulation. Data marketplace owners have terms and conditions that state how participants have to trade data on their platform: *"everyone comes here to do business and we have clear terms and conditions that say how to trade data"* (DM4). Another interviewee explained: *"smart contracts and other blockchain mechanisms are implemented. This way the data is secure"* (DM2). We started to recognize that data marketplace owners use contracts and formulate conditions to comply to regulation. The category data regulation is further enhanced with literature sources about contracts. Technical and non-technical sources provide information about bilateral and multilateral contractual relationships. For example, a TL source states that most data marketplaces sell data "via bilateral and negotiated contractual relationships" (Koutroumpis et al., 2017, p. 4). This triggered the connection between contracts and customer relationships. When we revised the interview data, the concept of contracts as means for data marketplace owners to maintain customer relationships further developed. One interviewee stated: *"at the moment when a customer purchases a data product from us, the conditions state what the data is and is not allowed to be used for"* (DM7). The control over data usage terms in contracts is explained in NTL3: "a data owner can sign a contract with a data user that forbids any distribution to or re-use by third parties" (Duch-Brown et al., 2017, p. 15). In the category customer relationships we explicate contracts as means to maintain trust and power relationships by the data marketplace owner and data sellers. Our enhanced categories and how we induce dimensions from these categories are further explained in section 4.1 – 4.3.

3.1.4 Aligning the Induced and Deduced Dimensions

Our preliminary taxonomy is based on the dimensions that we induce from the interviews with data marketplace owners and the dimensions that we deduce from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). In section 4.4 we align the derived dimensions and create our preliminary taxonomy. This provides an answer to our first sub-question *Based on an exploration of the business models of data marketplace owners in the B2B automotive industry and based on the existing generic business model taxonomies for data marketplaces, what dimensions can be derived to include in our preliminary business model taxonomy?* We further refine our preliminary taxonomy with business model characteristics derived from data marketplace cases in subsequent steps.

3.2 Online Desk Research

Online desk research is part of the second step in our taxonomy development process (see Figure 7). The goal in this step is to gather case documents from a selection of data marketplaces. Cases were selected based on the following selection criteria:

- The data marketplaces fit the definition of a data marketplace as described in section 1.2.1: a data marketplace matches buyers and sellers, facilitates transactions and provides an institutional infrastructure to trade machine-readable data
- The data marketplaces specialize in automotive data
- The data marketplaces are B2B
- Case documentation of the data marketplaces is in English

Through theoretical replication, based on orientation and ownership, we researched business models of a variety of cases that are spread over three cells of data marketplace types. Additionally, we selected two cases per cell for theoretical sampling. This results in six data marketplace cases from three data marketplace types. The data marketplace types are (i) data marketplaces with a hierarchical orientation and private ownership, (ii) data marketplaces with a mixed hierarchy and market orientation and consortium ownership and (iii) data marketplaces with a market orientation and independent ownership. We limit our analysis to a number of six data marketplaces to perform in-depth case analyses and create more specific business model insights than currently available in the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). To find cases from the data marketplace types in the B2B automotive industry we performed online desk research. We searched for data marketplaces that are past the conceptual stage. Due to the low number of data marketplaces in practice, there were few cases to choose from. It must be noted that IOTA and Ocean Protocol are data marketplaces which are still in the conceptual stage and trade data across industries. Data marketplaces with a market orientation and independent ownership that trade in automotive data and are past the conceptual stage were not found. We selected Ocean Protocol, because they are in the beta phase, almost ready for final release. IOTA is included due to their high number of 70 signed up participants. The data marketplaces are described in Table 10

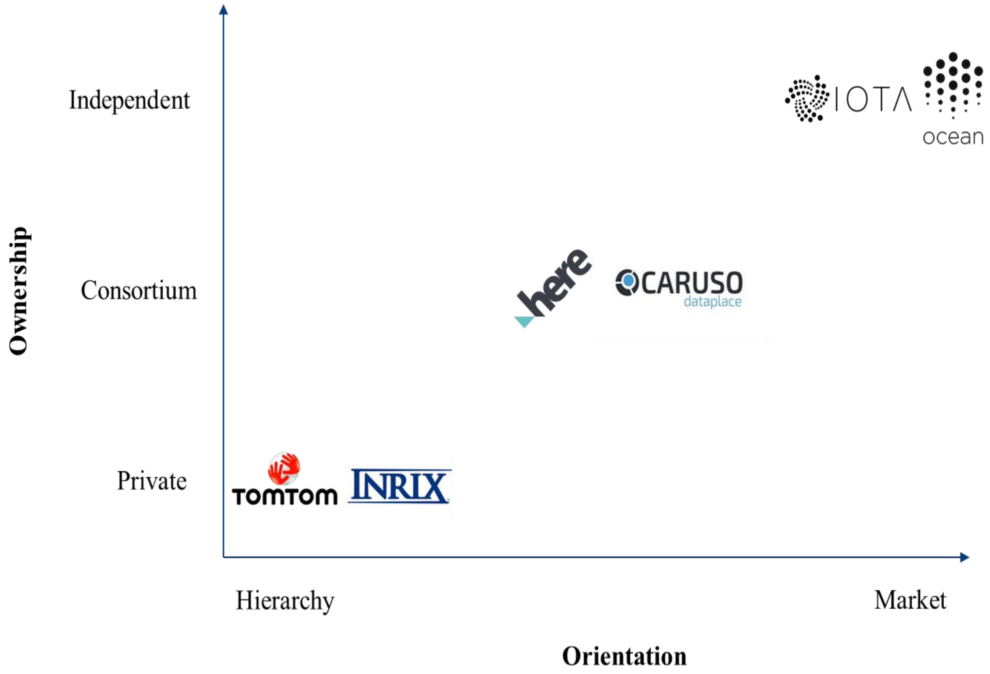

Figure 9: Selected data marketplace types

Table 10: Case descriptions

Data marketplace	Founded	Description
TomTom (TT) Hierarchical, private	1991	TomTom is a privately owned company that uses location technology to sell mapped data. They trade data in a hierarchically oriented, bilateral market. TomTom is well known for their sale of navigation boxes to end consumers. However in this study, we focus on the B2B segment of TomTom that concerns data trade between TomTom and their commercial buyers and sellers.
INRIX (IN) Hierarchical, private	2005	INRIX is also a privately owned company and applies location analytics to make road transportation more intelligent. INRIX trades data bilaterally with their commercial data sellers and buyers and serves public organizations. In addition to trading data, INRIX performs research on subjects such as road congestion, commuting time and vehicle carbon emission. Their research branch is out of scope in this thesis, because their research reports are in textual format and are not considered machine-readable data.
HERE (HE) Hierarchical/market, consortium	2015	HERE is formerly known as Navteq and was owned by Nokia. In 2015 the company was acquired by a consortium. Audi, BMW and Daimler are the main shareholders. HERE applies location technology to improve connected driving experiences. The HERE data marketplace has open access for any data seller, data buyer and third-party service provider to foster collaboration and share data among participants.
Caruso (CR) Hierarchical/market, consortium	2017	Caruso is founded by TecAlliance, a provider of vehicle data in the automotive industry. Besides TecAlliance, multinationals such as Bosch and Continental are shareholders of Caruso. The data marketplace is closed and only the consortium members and partners are allowed to trade at the data marketplace.
IOTA Market, independent	2017	IOTA is founded by the non-profit IOTA Foundation. IOTA focuses on the IoT market with the goal to enable secure data transactions between data sellers and buyers. The IOTA data marketplace has open access that allows many participants to trade data. IOTA is currently in the proof of concept phase.
Ocean Protocol (OP) Market, independent	2017	Ocean Protocol is a non-profit organization based in Singapore. Their data marketplace has open access to create an environment in which many data sellers and buyers can exploit data. The data marketplace is currently in its beta stage and is planned for a new release in Q3 of 2020. Ocean Protocol is particularly focused on AI. With high data volumes and trained algorithms they aim to advance AI development.

Information about the cases was collected from a variety of sources. The webpages, whitepapers and terms of use documents contain the most important information from the point of view of the data marketplace owner. We analyzed these sources first to get an impression of the vision and activities of the data marketplace owner. Additionally, reports or articles from external sources were consulted for an external view on the business model of the data marketplaces. Many online external news articles on the selected data marketplaces exist. For example, a Google search term stating “TomTom news article” generates over 2 million results. Within the limited time frame of this thesis we could not sort and analyze all these results. Therefore, we selected one external source as main source to provide us with information about the business model from an external perspective. To this end, Forbes is included as external source. Forbes is a renowned company, focusing on business, investing, technology, entrepreneurship, leadership and lifestyle. Only Caruso is not covered in any of the Forbes articles. Therefore, a report from Automat who performed an extensive market analysis on its competitors is included as an external source for Caruso. As additional external source, we included a Harvard case study for INRIX. This case study contains detailed information about INRIX’s business model. If we required more information about the business models after we analyzed these sources, news releases of the cases are included to reach saturation. An overview of the selected sources is presented in Table 11. A more detailed reference list of the case sources is included in Table 31 in Appendix E.

Table 11: Case sources

Sources	TomTom	INRIX	HERE	Caruso	IOTA	Ocean Protocol
Website	Main website develop portal	Main website	Main website Develop portal	Main website	Marketplace platform	Main website
Terms and conditions	Buyer Supplier	Site terms	Service terms	Privacy	Privacy	Privacy
Whitepaper	Product Annual report	Product	Product	Slides Live presentation	Technical	Technical Business Marketplace
External	Forbes articles	Forbes articles Harvard Business Review	Forbes articles	Automat report	Forbes articles	Forbes articles
Total	*9*	*8*	*7*	*5*	*5*	*7*

3.3 Content Analysis

By applying content analysis we completed the second step in our taxonomy development process (see Figure 7). The goal in this step is to specify the business model characteristics of the selected cases to answer the second sub-question: *What business model characteristics do owners of current data marketplaces in the B2B automotive industry apply?* Figure 10 shows the input-output diagram for this step. The preliminary taxonomy, that we created by applying the Grounded Theory method, and the case documentation, gathered from online desk research, form the input for content analysis. Our refined taxonomy is based on the business model dimensions and characteristics of the data marketplace cases that we selected.

Figure 10: Input-output diagram content analysis

Content analysis was applied to extract information from data sources. Content analysis is a qualitative data analysis technique that is commonly applied analyze large volumes of texts and infer categories in a systematic and replicable manner (Krippendorff, 2018; Mayring, 2004). An advantage of content analysis is that texts can be interpreted within a certain context (Krippendorff, 2018). Our research is placed within the context of business models. More specifically, we distinguish categories based on the value creation, value delivery and value capture components of data marketplaces. The interpretation of the texts can be performed by human coders, who can interpret semantically complex texts, and the texts can be interpreted using computer coding, where reliability of intercoder agreements is high (Krippendorff, 2018). We applied human coding. Since theory on business models for data marketplaces is not fully developed yet, there is no final set of coding rules that could be applied by a computer. Thus, human coding is more suitable to generate categories, due to the ability to define codes based on semantic interpretation.

We followed the inductive category development process by Mayring (2004), shown in Figure 11. Mayring (2004) makes a clear distinction between a content-analysis model to induce concepts and a model to deduce concepts. Krippendorff (2018) does not make this distinction in his content analysis method. Since our goal is to induce concepts, we followed the inductive category development process by Mayring (2004). This process contains six steps: (i) state the objects that should be classified (ii) define the categories (iii) induce new categories (iv) revise categories and do a formative check of reliability (v) perform a final iteration through the case documents (vi) interpret the results. Steps 2-4 were performed iteratively to improve our induction of taxonomy dimensions and characteristics. These steps are further explained in section 3.3.1 – 3.3.3.

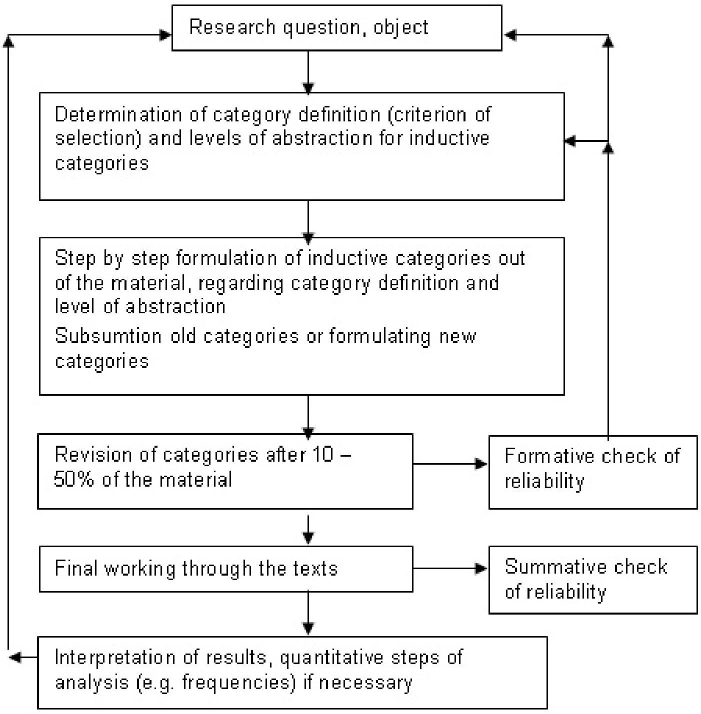

Figure 11: Inductive category development steps (Mayring, 2004)

3.3.1 Category Definition

In step 2, we defined the categories (see Figure 11). These categories are the business model dimensions that would be further refined in subsequent steps. Our preliminary taxonomy offered the initial description of these categories. These descriptions functioned as initial coding rules (see Table 12) to induce categories in step 3 of content-analysis. It must be noted that these are initial descriptions. The definitions of dimensions and characteristics are subject to change as iterations take place.

Table 12: Initial coding rules

If the unit of analysis describes, refers to or contains…	**Then assign the unit to node…**	**Then assign the node to node…**
the market in which the data marketplace is active	Domain	Customer segment
the data sellers and buyers who are matched at the data marketplace	Participants	Customer segment
the platform access terms for the customer	Platform access	Value proposition
the privacy protection of the stored data	Privacy	Value proposition
the source where the data on the data marketplace is collected from	Data source	Value proposition
the transformed state in which the data product is delivered	Data product	Value proposition
the quality guarantee of the purchased data	Data quality	Value proposition
the agreement that regulates data trade	Contract	Customer relationship
the communication, distribution and sales channels through which the value proposition is delivered	Channels	Channels
the storage location of data	Platform infrastructure	Key resources
the activities performed by the data marketplace owner that increase the value of the data	Data processing activities	Key activities
the alliances created to optimize the business model, reduce risk or acquire resources	Partners	Key partners

the way in which the data marketplace owner generates turnover by charging fees to its customers for a data transaction, marketplace membership, listing of data product, storage space or use of services	Revenue streams	Revenue
the pricing mechanism that is used to establish the price of the data output that is traded between the trading entities	Data pricing mechanism	Pricing model
the currency in which the payment is transferred	Payment currency	Pricing model

3.3.2 New Category Induction

In step 3, we induced categories from the case documentation (see Figure 11). Atlas.ti was used for our coding process. This is a qualitative data analysis software that is a helpful tool for especially large texts, visuals or audio data (Smit, 2002). Texts were labeled and highlighted in Atals.ti which resulted in codes. The ability to order, structure, retrieve and visualize information is a strength of Atlas.ti (Smit, 2002). Similar to the coding process explained in section 3.1.2, we applied initial and focused coding to induce categories. For initial coding, we applied line-by-line coding.

For example, a line on the website of TomTom states *"We license maps, navigation software and online services as components for applications, offering tailor-made solutions to meet customer's specific needs"* (TT-1). We assigned this line the following codes: *license products* and *offer tailor-made solutions*. Other lines from a TomTom whitepaper state *"For filtering, an enhanced data analysis is necessary, for example to separate handsets that are used in a train. As a typical speed pattern appears when calls are coming from trains, because all handsets have the same speed and handover events, these data can be taken out"* (TT-3a). We assigned the following codes to these lines: *filter data*, *analyze data*, *separate data* and *remove data*. Similar to these examples, initial codes for all documents were generated per case. This results in almost 1000 initial codes in total (see appendix E.1–E.6).

After initial coding was finished for each case, we performed focused coding to categorize the codes. This was accomplished by applying axial coding. Axial coding is a technique to link categories and subcategories (Charmaz, 2006; Strauss & Corbin, 1998). This technique fits our objective to connect taxonomy dimensions to characteristics.

Take the example of the initial codes from the TomTom documents as described above. TomTom licenses products and offers these as tailor-made solutions. This goes beyond selling a product. One of the core products of TomTom are their maps. The maps are based on analyzed location data. Other initial codes are *payment service*, *deliver product* and *customized supply response*. These codes do not fit our description of the dimension *data product* or *data processing activities* as formulated in the initial coding rules in Table 12. A new dimension emerged that we did not consider before; data service. We did not replace the dimensions data product or data processing activities with data service, but added this dimension as a new one. The creation of a new dimension needed to be revised, which is explained in section 3.3.3. The initial codes *filter data, analyze data, separate data* and *remove data* fit our coding rule of *data processing activities* (see Table 12). Therefore, data filtering, analyzing, separation and removal are characteristics of the dimension data processing activities.

3.3.3 Category Revision

In Step 4 of the inductive category development approach, the existing and newly created categories were revised. Mayring (2004) advises to revise the categories after 10-50% of the material is analyzed as a formative check of reliability. During these revisions, categories were renamed and codes were removed or added. Because codes were created in high quantities, we preferred to revise categories in rather smaller than larger iterations. However, if categories are revised based on only 1 case, our judgement to distinct one dimension or characteristic from another could not be trusted. Especially not at early stages when codes from one case could not be compared to another. Therefore, we revised categories in steps of 2 cases out of 6. This means that revision took place after 33% of the material was analyzed.

The taxonomy development iterations are shown in Table 13. The first column, iterations 1-3, shows the dimensions that were derived by applying the grounded theory method. These dimension formed the input for our content analysis. Iterations 4-7 are the steps that were performed during content analysis. The new or adapted dimensions are shown in blue and removed dimensions are blocked gray. During iteration 4, we coded documents of TomTom and INRIX. In section 3.3.2 an example of the induction of a new category with the codes *license products*, *offer tailor-made solutions*, *payment service*, *deliver product* and *customized supply response* is described. Initially, we only considered the dimension *data product* in the value proposition of a data marketplace. During iteration 4, this dimension was split into two new dimensions: *data output* and *data service*. We categorized the codes *license products, offer tailor-made solutions*, *payment service*, *deliver product* and *customized supply response* in the characteristic *customized map service* of the dimension *data service*. During iteration 5, we changed the naming of dimension *data source* into *data input*. During iteration 6, the documents of IOTA and Ocean Protocol were analyzed. During this iteration, we removed the dimension *data input* to avoid overlap with the dimension *participants*. Data sellers are participants at the data marketplaces and provide data input for the data marketplace. The characteristics that we defined did not make a clear distinction between the two dimensions. Therefore, we removed one of the dimensions. After all documents were coded, we performed a summative check of reliability in iteration 7. This step is required to revise whether all codes are consistent and representative for the texts that are coded (Mayring, 2004). No new dimensions or characteristics should be generated during this revision as stated in one of the objective ending conditions in Table 1. A complete overview of the taxonomy development iterations is included in appendix F and the satisfaction of the ending conditions per iteration is included in appendix I. We describe the business dimensions and characteristics of our refined taxonomy in chapter 5.

Table 13: Iterations for inductive category development

Step	*Iteration 1-3*	*Iteration 4*	*Iteration 5*	*Iteration 6*	*Iteration 7*
Approach	Grounded Theory	Empirical-to-conceptual	Empirical-to-conceptual	Empirical-to-conceptual	Empirical-to-conceptual
Value creation	Domain	Domain	Domain	Domain	Domain
	Participants	Participants	Participants	Participants	Participants
	Privacy	Privacy	Privacy	Privacy	Privacy
	Data source	Data source	Data input	Data input	
	Data product	Data service	Data service	Data service	Data service
		Data output	Data output	Data output	Data output
	Data quality	Data quality	Data quality	Data quality	Data quality
	Contract	Contract	Contract	Contract	Contract
Value delivery	Platform access	Platform access	Platform access	Platform access	Platform access
	Platform infrastructure	Platform infrastructure	Platform infrastructure	Platform infrastructure	Platform infrastructure
	Data processing activities	Data processing activities	Data processing activities	Data processing activities	Data processing activities
	Partnerships	Partnerships	Partnerships	Partnerships	
Value capture	Revenue streams	Revenue streams	Revenue streams	Revenue streams	Revenue streams
	Data pricing mechanism	Data pricing mechanism	Data pricing mechanism	Data pricing mechanism	Data pricing mechanism
	Payment currency	Payment currency	Payment currency	Payment currency	Payment currency

New dimension

Removed dimension

3.4 Cross-Case Analysis

Performing a cross-case analysis was the third step in our taxonomy development process (see Figure 7). Our goal in this step is to recognize and interpret patterns in our taxonomy and answer the third sub-question: *What business model archetypes can be identified for data marketplaces?* Figure 12 shows the input-output diagram for this step. We generated business model patterns from our refined taxonomy and created business model archetypes by performing a cross-case analysis

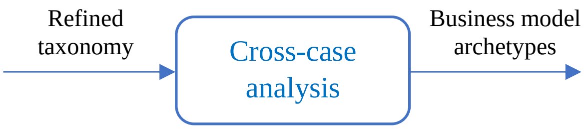

Figure 12: Input-output diagram cross-case analysis

A way to demonstrate the findings of taxonomies is to present the patterns as business model archetypes. Archetypes are reoccurring patterns in the combinations of taxonomy characteristics (Oberländer et al., 2019). The recognition of patterns goes beyond the taxonomy development method as described by Nickerson et al. (2013). Oberländer et al. (2019) identify two main methods to recognize patterns in taxonomies. These are the quantitative clustering analysis and the case study method.

A quantitative clustering analysis computes the characteristics into clusters, as applied by Hodapp et al. (2019), Oberländer et al. (2019) and Täuscher & Laudien (2017). Quantitative clustering analysis can be an objective method to observe patterns and generate business model archetypes. However, it is not a suitable method in our research, because we did not classify enough cases in the taxonomy for a quantitative clustering analysis. Fruhwirth et al. (2020) recognize that quantitative clustering is more suitable for taxonomies in which a high number of objects are classified. Their taxonomy counts a classification of only 20 objects, which is why they chose to identify patterns using the case study method. This is a suitable method to compare cases. Eisenhardt (1989) names three tactics to compare cases. First, within-group similarities and intergroup differences can be researched per dimension. Based on the within-group similarities and across-group differences, patterns among the cases emerge that lead to business model archetypes. This is a thorough way to generate patterns but can lead to a large amount of comparisons when the taxonomy includes many dimensions. Therefore, this is not the suitable tactic to apply within the timeframe of this research. Second, Eisenhardt (1989) suggests to select and compare pairs of cases. By searching for differences between seemingly similar cases, presumptions can be broken and a more sophisticated understanding of the patterns emerges. Third, the cases can be compared based on data sources. Four different source types are included per case (see Table 11) with multiple sources belonging to one type. Patterns from data sources contribute to unique insights and are corroborated with evidence from similar sources of all cases (Eisenhardt, 1989). However, comparison of the cases by source is a tactic that leads to highly granular comparisons. Because the information is already aggregated in the taxonomy characteristics, we prefer a higher-level case comparison over the comparison by source. The pairwise comparison tactic allows for a comparison of business model characteristics between cases. We applied this tactic to identify patterns and derive business model archetypes.

Based on the outcome of our refined taxonomy, we paired seemingly similar cases. In our pairwise comparison we aimed to go beyond initial impressions and break simplistic frames. If required, groupings were adjusted by the end of the analysis. Our pairwise comparison and business model archetypes are presented in chapter 6.

3.5 Semi-Structured Expert Interviews

In this section, the final step in our taxonomy development process is explained (see Figure 7). The goal in this step is to evaluate our taxonomy and answer the fourth sub-question: *Is our business model taxonomy evaluated as complete and useful?* Figure 13 shows the input-output diagram for this step.

Figure 13: Input-output diagram expert interviews

Evaluation of our taxonomy is key to ensure the taxonomy is complete and useful. However, the evaluation of taxonomies is often overlooked (Szopinski et al., 2019). Most researchers iteratively evaluate their taxonomy throughout the development process (Oberländer et al., 2019). For example, Fruhwirth et al. (2020) iteratively evaluated whether their taxonomy met all ending conditions during the development process. Similarly, we iteratively assessed the objective ending conditions as we were building our taxonomy. Furthermore, taxonomies can be evaluated after they are completed. Nickerson et al. (2013) suggest a taxonomy should be evaluated based on how others use it. However, a clear method to evaluate the completeness and usefulness of a taxonomy when the development process is finished is still missing. Therefore, Szopinski et al. (2019) propose a number of evaluation methods based on an extensive literature review of 196 studies in which taxonomies are developed. This results in 11 methodologies for evaluation. We opted for the method of expert interviews, because of two main reasons:

- Expert interviews are suitable to decide on the completeness of our taxonomy. The possible ambiguous interpretation of business model components of data marketplaces requires multiple people to evaluate the taxonomy. Various perspectives on the business model components from experts in data marketplaces contribute to identifying new dimensions or justifying the taxonomy as complete.
- Expert interviews are suitable to decide on the usefulness of our taxonomy. Nickerson et al. (2013) and Szopinski et al. (2019) state that a taxonomy can be used by researchers as well as practitioners, but do not elaborate when a taxonomy is considered useful. Nickerson et al. (2013) recognize that assessing a taxonomy as useful is difficult and that it has to be observed. They speculate about the use of a taxonomy and suggest to establish the usefulness in concordance with the intended users. This can be achieved through expert interviews. The interviewees can indicate how they envision the usage of the taxonomy and whether the taxonomy meets that purpose.

Our taxonomy is evaluated with Markus Spiekermann and Michael Fruhwirth. They are two researchers who designed the available business model taxonomies for data marketplaces that we built upon. We consider Spiekermann and Fruhwirth experts in both the field of data marketplaces and business models. Spiekermann (2019) developed a taxonomy to compare existing data marketplaces that "presents the essential elements of the business model in the form of a morphological box and makes the existing solutions comparable on the basis of defined dimensions" (p. 3). Fruhwirth et al. (2020) developed a taxonomy to "provide an overview of the current business models of data marketplaces" (p. 5744). The purpose of our taxonomy is to *identify the business models that data marketplaces in the B2B automotive industry implement*. We further evaluated this purpose with the experts to determine the completeness and usefulness of our taxonomy from their point of view. Although our taxonomy is intended to be used by researchers and practitioners, we only evaluated our taxonomy with researchers due to the limited timeframe of this thesis. We recommend taxonomy evaluation with practitioners for future research.

The interviews were semi-structured. We prefer the semi-structured interview method over the structured interview method because it enables interviewees to add new elements where necessary (Galletta, 2013). Szopinski et al. (2019) note that unstructured interviews can also be conducted to evaluate taxonomies as done by Herzfeldt et al. (2012). However, there are pre-defined conditions, such as the subjective ending conditions of taxonomy development, that we needed to evaluate. Therefore, an unstructured interview method is not suitable. Keller & König (2014) applied semi-structured interviews to evaluate their taxonomy as well, which proved as a suitable method to test the ending conditions.

The interviewees were sent information about our research and a consent form that is included in appendix G. Background information about the thesis and the evaluation steps were explained to them. Our evaluation process contains four steps (see Table 14). First, the selected data marketplaces types were explained. This concerns our choice to include data marketplaces with hierarchical and market orientation. During the second step, the taxonomy was discussed. This step started with evaluating the completeness during which the interviewee had time to look at the taxonomy and indicate whether relevant characteristics and dimensions were missing to describe a business model of a data marketplace. Subsequently, we tested the subjective ending conditions based on questions suggested by Nickerson et al. (2013, p. 344). To evaluate whether the taxonomy is concise, they propose to ask: “Does the number of dimensions allow the taxonomy to be meaningful without being unwieldy or overwhelming?” However, “meaningful” and “overwhelming” are ambiguous phrases. Therefore, we asked the interviewees whether dimensions or characteristics should be removed from our taxonomy. If the interviewees could not identify dimensions and characteristics that should be removed, we perceived this as an indication for a concise taxonomy. For comprehensiveness, Nickerson et al. (2013, p. 344) suggest to ask: “Can all objects or a (random) sample of objects within the domain of interest be classified?”. This is a question that is hard to answer for the interviewees, because they are not expected to have complete knowledge over the sample of data marketplace cases in this research. Therefore, we asked them whether the data marketplace DAWEX could be classified in our taxonomy. DAWEX is a data marketplace that we did not include in our research but both Spiekermann (2019) and Fruhwirth et al. (2020) did. If they indicated that DAWEX can be classified in our taxonomy we noted that as an indication for a comprehensive taxonomy. Next, the usefulness of our taxonomy was explored. The interviewees were asked how they envision the use of our taxonomy and whether our taxonomy is useful for that purpose. During the third evaluation step, our archetypes were discussed. Similar to the usefulness of our taxonomy, we explored the usefulness of our archetypes with the interviewees. Because only two expert interviews were conducted, we did not require a structured analysis method to retrieve information from the interview data. Feedback was retrieved from the transcribed interviews. The evaluation results are discussed in chapter 7.

Table 14: Semi-structured interview questions

Step	Evaluation criteria	Interview question
Evaluate the data marketplace sample	Completeness	Are relevant data marketplaces missing to form a representative sample for B2B data marketplaces in the automotive industry?
	Completeness	Are characteristics missing in the taxonomy to describe a respective dimension? Are dimensions missing in the taxonomy to describe a business model of a data marketplace?
	Concise	Should dimensions or characteristics be removed from the taxonomy for the benefit of overview?
	Robust	Do the dimensions and characteristics provide for differentiation among objects sufficient to be of interest?

Evaluate the taxonomy	Comprehensive	Can the business model of DAWEX be classified in this taxonomy?
	Extendible	Can a new dimension or a new characteristic of an existing dimension be easily added?
	Explanatory	What do the dimensions and characteristic explain about an object?
	Usefulness	How do you envision the usage of this taxonomy? Do you determine the taxonomy useful for that purpose?
Evaluate the archetypes	Usefulness	How do you envision the usage of the archetypes? Do you determine the archetypes useful for that purpose?
Conclude the evaluation	-	Do you have any additional remarks?

3.6 Methodological Limitations

The methodologies that we applied face limitations that influence the quality of the research results. These limitations must be considered in discussing the results.

First, following the grounded theory method creates the risk of being influenced by theories one is familiar with before starting the explorative research. The goal of grounded theory is to establish theory without any prejudice about the desired research result (Charmaz, 2006). However, in this research the Business Model Canvas (Osterwalder & Pigneur, 2010) and STOF model (Bouwman et al., 2008) form the first set of business model components that are considered for our taxonomy. This could bias us in formulating the components and could result in overlooking new business model components or important relationships for data marketplaces.

Second, we used limited resources during content analysis. Six cases were selected from the automotive industry. Our results are based on those cases, which goes at expense of external validity. Whether our results are applicable to data marketplaces in other industries was not tested. However, we reached internal validity by spreading the selected data marketplaces over different orientation and ownership types. The orientation and ownership of data marketplaces are not sector specific and can still generate results for a more generalizable audience. Furthermore, due to time constraints we selected a limited number of case sources. Case documentation that we retrieved from the websites of the respective data marketplace are overrepresented in the case sources. This causes bias in the information derived from these sources. This bias is compensated by the inclusion of the external sources from Forbes. The risk of missing business model components is reduced by reaching theoretical saturation in the content-analysis. Theoretical saturation is reached when no new dimensions or characteristics emerge (Charmaz, 2006). We only ended our taxonomy development process when "no new dimensions or characteristics were added in the last iteration" (see OE4, Table 1). This should be the case in the fifth step of the inductive category development by Mayring (2004) (see section 3.4).

Third, the selection of pairwise comparison for the cross-case analysis can lead to bias in our business model archetypes. We compared cases in groups of two, based on assumed similarity between cases. Looking for juxtapositions in the groupings should lower the risk of biased archetypes, but does not eliminate this risk. An analysis of within-group similarities coupled with intergroup difference for all dimensions of all cases in the taxonomy would create a more complete comparative analysis. However, this results in a high number of comparisons (the number of dimensions multiplied by the number of cases). Therefore, we performed the pairwise comparison analysis with the risk of forming biased archetypes.

Fourth, the expert interviews are limited in variation of the interviewees. This is reflected in the number of conducted interviews and the experience of the interviewees. We evaluated our taxonomy with 2 experts. That is the minimum. An evaluation process with more experts would benefit the validity of our results. We selected experts based on their experience in the development of a business model taxonomy for data marketplaces. The interviewees can be biased towards their own choices in taxonomy dimensions and characteristics. No practitioners such as data marketplace owners were interviewed to evaluate our taxonomy. This is advised for future research.

Finally, all methodologies are limited by the fact that they are performed by one researcher. In qualitative research and especially in case analyses it is advised to work in multidisciplinary groups. Different perspectives in qualitative research contribute to the creation of new ideas (Eisenhardt, 1989). Furthermore, confidence in the results increase when multiple researchers can agree on the outcome (Eisenhardt, 1989). Conflicting perceptions also contribute to higher trustworthiness of the results as they protect against premature conclusions. Despite the performance of the research by one researcher, we had various reflection moments to increase the trustworthiness of our results. First of all, weekly discussions with a PhD student from the University of St. Gallen were held about our research progress. During these discussions the PhD student indicated where our reasoning was unclear or incorrect and needed to be improved. We also discussed the results with two supervisors from the TU Delft who gave detailed feedback in the middle of the research process and towards the end. Moreover, we performed eight iterations to develop our taxonomy. In each iteration we checked whether codes were consistent and whether the dimensions and characteristics in our taxonomy needed to be adapted based on new data that we gathered from interviews, technical literature and non-technical literature. These reflections increase the trustworthiness of our results.

4. Taxonomy Conceptualization

A business model taxonomy for data marketplaces shows the essential business model dimensions that enable comparison of cases (Spiekermann, 2019). In this chapter we explore business model categories to derive business model dimensions for our preliminary taxonomy. As explained in section 3.1 *data regulation*, *customer relationships*, *platform infrastructure*, *data processing activities* and *revenue models* were main topics of discussion during interviews with data marketplace owners. We derive business model dimensions from these categories to distinguish how data marketplace owners create, deliver and capture value.

This chapter is structured as follows. In section 4.1, the categories data regulation and customer relationships are explained. We derive the dimension *contract* from these categories, part of the meta-characteristic *value creation* in our taxonomy. In section 4.2, the categories *platform infrastructure* and *data processing activities* are explained. These categories lead to identical business model dimensions, part of the meta-characteristic *value delivery*. In section 4.3, the category revenue model is explored. We derive the dimensions *revenue streams* and *data pricing mechanism* from this category. These dimensions are part of the meta-characteristic *value capture* in our taxonomy. Next, the induced dimensions are aligned with the dimensions that we deduce from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) in section 4.4. This results in our preliminary taxonomy. Finally, an answer to the first sub-question is provided in section 4.5.

4.1 Data Regulation and Customer Relationships Leading to the Dimension Contract

Data marketplace owners have to regulate data trade at their platform to comply with data regulation. Data regulation was often mentioned as a challenge by the interviewees, because regulatory guidelines leave room for interpretation. Data marketplace owners incorporate data privacy and data processing rights into their contracts to adhere to privacy regulation. This is explained in section 4.1.1. Data regulation is related to the creation of trust and power relationships between data marketplaces and their customers, which is explained in section 4.1.2. In section 4.1.3 we conclude with deriving the dimension *contract* as part of the meta-characteristic *value creation* in our taxonomy.

4.1.1 The Regulation of Data Trade in a Contract

The distinctive nature of data goods compared to physical goods requires special attention to data regulation. When someone uses a physical object, another person can be excluded from using it. However, data goods can be used by many people at the same time without excluding anyone. The lack of excludability raises issues concerning privacy and ownership (Duch-Brown et al., 2017). How is data privacy preserved? Who is responsible for safeguarding privacy? Who is allowed to access data? What purposes may the data be used for? These are questions that data marketplace owners need to consider in their business model. These questions are addressed in data regulation. The General Data Protection Regulation (GDPR), contract law, EU Database Directive and Trade Secrets Protection Directive contain rules that data marketplace owners need to comply with. An overview of these regulations is provided in Table 15 with key insights for data marketplaces.

Table 15: Data regulation

Regulation	Key insights
GDPR (2016)	• Data subjects have the right to rectify, be forgotten and restrict processing of their personal data (chapter 3, article 16-18) • Data controllers must have written consent to process the data, processing activities must be recorded, measures to secure processing activities must be taken (article 28-32) • Few guidelines are given concerning ownership rights.
Contract law	• Contract law is defined per nation • Contract law follows the Rome Convention for the Protection of Performers, Producers of Phonograms, and Broadcasting organizations (1961). • The rights and obligations of sellers and buyers are defined in contract law

EU Database Directive (1996)	• Protects data in electronic databases, not the insights extracted from data.
Trade Secrets Protection Directive (2016)	• Protects information that is not generally known or accessible by the public • Defines the right to be compensated for damages caused by dishonest practices

Data marketplace owners need to keep up with regulatory changes and have to ensure that mechanisms in their business model are within the boundaries of the institutional framework. Most interviewees acknowledged data regulation as a challenge, because rules are not always clearly defined and leave room for interpretation. However, the data marketplace owners explained that they incorporate rules to preserve data privacy and adhere to regulation. One data marketplace owner explained that they protect data privacy by anonymizing data that is stored: *"stored data is anonymized in such a way that one cannot derive what car would drive to what address and what individual the information belongs to"* (DM6). Another data marketplace owner explained that the data they trade, complies with GDPR and that they do not store privacy sensitive information about their users: *"all data items that are put in the data marketplace comply with GDPR. We do not store any user data"* (DM2). In addition to storing data in a secure manner, data marketplace owners incorporate data regulation into their contracts. For example, data marketplace owners agree with their data sellers on the data that will be traded at the data marketplace. Consent is given to trade data, as an interviewee explained: *"we function as a consent management hub. We facilitate communication between a newly developed application and an Original Equipment Manufacturer to give consent to use parameters of a car"* (DM6). The rules to trade data at the data marketplaces are clarified in their terms and conditions: *"everyone comes here to do business and we have clear terms and conditions that say how to trade data"* (DM4). The contract conditions regulate how actors may trade data.

Contracts are the institutional infrastructure that data marketplaces provide. In the contracts, data marketplace owners define what data is offered and to whom. Ramirez et al. (2014) explain that contracts clarify whether the data marketplace acquires ownership over the data, for what purpose the data may be used and what the rights are to resell the data. In addition, a description of the data, the method used to transfer the data and the data update frequency are included in the contract. In short, contracts are means to clarify data ownership and mitigate privacy issues of data marketplaces. Truong et al. (2012) list five main properties of data contracts. These are (i) data rights that define whether the data buyer is authorized to for example alter and reproduce the data (ii) the data quality such as accuracy and consistency expectations (iii) regulatory compliance to laws such as the GDPR, (iv) the pricing model that is often transaction or subscription based and (v) control and relationship terms.

Two main contract types are implemented in different data marketplaces. Data marketplaces with a hierarchical orientation have negotiated contracts and data marketplaces with a market orientation have standardized contracts. At data marketplaces with a hierarchical orientation, participants often trade data bilaterally (Koutroumpis et al., 2017). Most contracts are bilateral, because this enables data sellers to select the actors who may access their data and prohibit data buyers to spread the data to third parties (Duch-Brown et al., 2017). The contract conditions are negotiated between the trading parties and contain clauses that the data marketplace owner specifies according to the demands of the data seller. However, when the number of customers at a data marketplace increases, negotiated contracts become cost intensive for the data marketplace owner because transaction costs increase. Therefore, data marketplaces with a market orientation have standardized contract conditions (Koutroumpis et al., 2017). Smart contracts are an example of standardized agreements in decentralized data marketplaces. Such contracts run on distributed ledger technology (DLT) (Ølnes et al., 2017). Data marketplace owners who implement smart contracts promote safe transactions, as there is no intermediary who can manipulate the data. The data seller and data buyer have to give consent to the terms in the standardized contract. After consent is given, the transaction is performed automatically (Savelyev, 2017). This enables fast transactions when a high number of trading parties is involved.

Overall, the advantage of negotiated contracts is the ability for data sellers to choose ownership and usage rights. However, the disadvantage is that the transaction costs go up when the number of trading parties increases. This is solved by data marketplace owners who apply standardized contracts with standardized trading conditions. The automatization of the transaction process in standardized contracts enables lower operational costs and higher transaction speed than is the case with negotiated contracts. The elimination of a third party intermediary causes higher transparency in data processing. Figure 14 shows the implementation of the contracts in data marketplaces with hierarchical orientation and market orientation. In hierarchical structures, negotiated contracts allow bilateral exchange. In market structures, standardized contracts enable multilateral exchange.

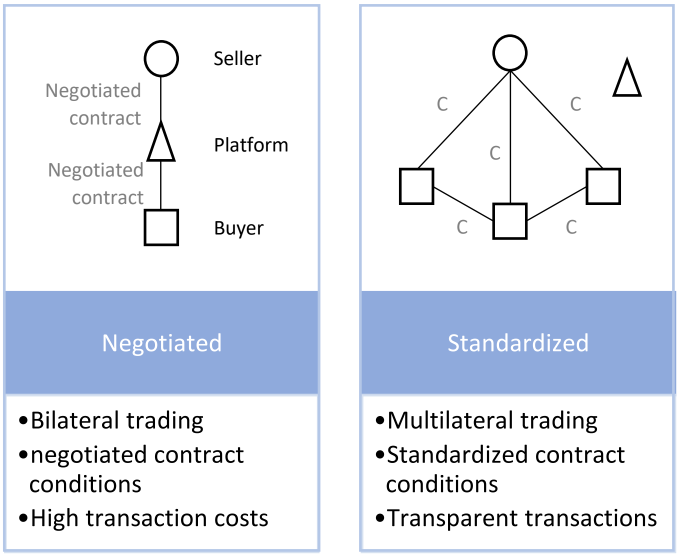

Figure 14: Contract types

4.1.2 The Establishment of Trust and Power Relationships in a Contract

Trust and power relationships influence enterprises in their decision to enter a marketplace (Christiaanse & Markus, 2002; Nicolaou & McKnight, 2006; Pavlou, 2002). Data marketplace owners establish these relationships using contracts.

Data marketplace owners establish trust with their participants when they can ensure safe operations (Koutroumpis et al., 2017). For the data marketplace owner, this entails the implementation of data protection, security mechanisms and acting conform privacy regulation. Pavlou (2002) calls this institutional trust. Institutional trust is created by transparently implementing licenses, contracts and regulations. A data marketplace owner (DM1) explained that they maintain strict privacy rules that are communicated to the customer. For example, if data buyers request data that violates privacy regulation, the data marketplace owner denies this request. Another data marketplace owner explained that they are transparent towards data buyers about the usage terms in their contract: *"the usage terms define whether the data seller imposes copyrights for certain regions"* (DM7). Both these examples constitute institutional trust that the data marketplace owner creates by means of transparent usage conditions in their contracts.

The playing field in the automotive industry is dominated by power relationships (Martens & Mueller-langer, 2018). According to Martens & Mueller-Langer (2018), OEMs have monopoly power in the automotive market, because they design the data architecture of cars in such a way that they retain exclusive control over the data. The OEMs are not eager to participate at a data marketplace because they fear to lose exclusive access to their data. To stimulate participation of OEMs and preserve power at the data seller, some data marketplace owners incorporate certain usage conditions in their negotiated contracts with OEMs. One of the interviewees explained that their data sellers can demand usage conditions in their contracts that restrict certain participants from accessing the data, stating *"certain* *sources can set specific conditions for the use of data for applications. Our data suppliers can exclude data usage for specific applications, regions, type of vehicle or end user"* (DM7). The OEMs can exert power through negotiated contract conditions. According to another interviewee, OEMs should not have the exclusive control over their data access: *"OEMs are starting to realize that they cannot control data access anymore. They are the guardian of the data that is owned by the person who produced the data, the driver of a vehicle"* (DM6). This data marketplace owner negotiates with OEMs to provide third party service providers access to data from the OEM. Third party service providers, such as telecom providers, insurance companies and navigation companies, rely on the data provided by OEMs to develop their services (Martens & Mueller-langer, 2018). One interviewee thinks that without government control, powerful stakeholders in the automotive industry may not share their data at an open data marketplace. He implied that companies can be forced to make their data available at data marketplaces (national data access points) via new regulation: *"With newer regulations, data needs to be made available to national access points in Europe. […] We are discussing how to better reach international organizations with other European member states to make sure that data provision to the access points is known and is considered in regards to the EU law"* (DM3). The European Commission (EC) (2020) announced that they want to create a European data space, a single market, to trade data across industries in Europe. The EC will create a regulatory framework that defines data access and usage rules that are "fair, practical and clear" for a single market (European Commission, 2020, p. 5). Their strategy is to stimulate voluntary data sharing between companies. However, when large players unilaterally restrict data access at the marketplace to preserve their own power advantage, the EC could consider stricter regulation to stimulate the openness and fairness of the market. This is not practiced yet.

4.1.3 Contract as a Business Model Dimension, Part of Value Creation

Although data regulation may be identified as a challenge, we noted that data marketplace owners are able to interpret regulation and incorporate rules into their contracts. By transparently communicating rules about for example data privacy, data marketplace owners create institutional trust with their customers. Furthermore, the data seller and data marketplace owner agree about data processing and distribution in negotiated contract conditions. Data sellers can demand the data marketplace owner or data buyer to include certain usage conditions in the contract to retain power over their data. This is different from standardized contracts in which the same conditions apply to everyone. Standardized contracts have lower transactions costs than negotiated contracts. This makes them suitable for multilateral data trade. Overall, data marketplace owners apply *negotiated* or *standardized* contracts to adhere to data regulation and to establish customer relationships. Data regulation is an external factor and is not part of the business model of data marketplaces. Therefore, we include *contract* as a dimension of the component *customer relationships* in our taxonomy, part of the meta-characteristic *value creation*.

4.2 Platform Infrastructure and Data Processing Activities Leading to Identical Dimensions

Data marketplace owners use their platform infrastructure to store data. They store data in a centralized or decentralized location as we explain in section 4.2.1. The storage location impacts the data processing activities that a data marketplace owner can perform. In section 4.2.2 we introduce these activities. Data marketplace owners can either perform all of the processing activities or a limited number of those activities. This impacts the transformed state of the data that is traded on the data marketplace. Finally, in section 4.2.3, we conclude by deriving the identical dimensions *platform infrastructure* and *data processing activities* as part of the meta-characteristic *value delivery* in our taxonomy.

4.2.1 The Storage of Data in a Centralized or Decentralized Platform Infrastructure

The platform infrastructure indicates whether data marketplace owners deploy a centralized or decentralized platform architecture. A centralized approach enables data access and storage at a central location such as the cloud (Koutroumpis et al., 2017). The data marketplace owner gets ownership over the data that is stored in their centralized platform infrastructure and has the ability to restrict access to the platform (Koutroumpis et al., 2017). In a decentralized platform infrastructure, data sellers and buyers exchange data directly, without the interference of a central intermediary (Koutroumpis et al., 2020). Spiekermann (2019) explains that in a decentralized platform infrastructure, the data seller benefits from control over their data, but data processing and storage get more complicated for the data marketplace owner. In a decentralized platform infrastructure data is stored on the distributed ledger. The DLT keeps track of the history of all transactions and reports this to the buyers and sellers who trade data (Koutroumpis et al., 2017). The ability of data sellers and buyers to verify the transactions makes a decentralized platform infrastructure transparent. A data marketplace owner stated the following about data storage at their platform: *"we do not know where the data is. We only are the protocol in between that helps transactions happen"* (DM4). They deploy a decentralized platform infrastructure without a central point of control. They do not store or process data. A decentralized platform infrastructure increases data sovereignty for the data seller and buyer but complicates data storage and analysis for the data marketplace owner. Data storage and analysis are part of the data processing activities explained in section 4.2.2.

4.2.2 Data Processing Activities to Transform Data

The data processing activities performed by a data marketplace owner impact the value proposition of a data marketplace. Thomas & Leiponen (2016) explain that organizations can directly sell data little investment in data transformation or they can process and analyze data to deliver a data service. Essentially, all data marketplace owners perform at least two data processing activities; data collection and data distribution. They collect data from data sellers and distribute data to data buyers. In between those activities, data marketplace owners can perform varying processing activities. The data processing activities differentiate the data product or service for the customer. Allee (2008) describes three different forms of data value processing. These are (i) value adding: insights are added to data and traded as a new intangible good, (ii) value extension: the asset is made available to new partners and (iii) value conversion: an intangible asset is transferred into a tangible asset. In this section we define these processing activities on a more granular level, shown in Figure 15. The data processing activities are based on the big data value chain by Curry (2016) (acquisition, analysis, curation, storage and usage) and the data processing activities mentioned by Koutroumpis et al. (2017) (data import, storage, transformation, aggregation, analysis and delivery). In the following paragraphs these activities are further explained and elaborated with examples from the interviews with data marketplace owners.

Figure 15: Data processing activities

4.2.2.1 Data Collection

All data marketplace owners collect and distribute data. During data collection, data marketplace participants agree on the data that will be traded and processed. One data marketplace participant explains that the negotiation process to collect data is complex. Sometimes the process to acquire data takes years. The data marketplace performs data collection for their customers: *"The data buyers simply want to have the data. They do not want to negotiate for years to get the data. That is what our data marketplace facilitates"* (DM6). Another interviewee explained that during data collection, terms and conditions are agreed on: *"The data that we buy is always accompanied with terms and conditions. We try to standardize those as much as possible. However, there is a difference in terms and conditions among suppliers. So, the moment a customer buys one of our products, the agreement states under what conditions the data may or may not be used"* (DM7). The conditions that data marketplace owners, sellers and buyers agree upon during data collection, influence what data processing activities may be performed with the data, such as data enhancement, data analysis and data distribution. Before those activities are performed, data is standardized to enable easy exchange of data.

4.2.2.2 Data Standardization

Data standardization improves the interoperability of data. Data interoperability is one of the main tasks of data marketplace owners (Deichmann et al., 2016). During this task, data marketplace owners format data from various sources into one type, as one interviewee explained: *"…we facilitate IT integration. We enable standardization of the data in such a way that it results in one common language to easily deliver data to consumers"* (DM6). A single data format contributes to the unambiguous interpretation of data by all involved trading parties (Janssen & Verbraeck, 2005). Not all data marketplace owners perform data standardization. One interviewee explained that the data sellers who want to trade data at the marketplace need to standardize it themselves: *"if they [data sellers] want to put their data in our data marketplace, they need to divide the meta data, find the size, type and many other things"* (DM2). The data marketplace owner provides the format in which the data seller needs to organize their data: *"we give a tool to the users, so they can easily map the data sources. The tool incorporates the meta data in the marketplace"* (DM2). Once data is standardized, data can be tagged to improve the search process of data at the data marketplace (Koutroumpis et al., 2017). The tags are part of the meta data that data marketplace owners publish on their platform. The transformation to standardized data enables data buyers to find information that fulfills their needs.

4.2.2.3 Data Cleansing

Data cleansing contributes to the data quality at a data marketplace. During data cleansing, data marketplace owners check the data consistency and verify the data content. One interviewee explained that they clean data in collaboration with their customers: *"A challenge for most of our customers who make use of digital maps is to ensure that the provided data is correct. We collaborate with our customers to detect data that is incorrect and automatically improves this in the system. This brings us the advantage to improve the digital map without manual interaction"* (DM7). Other data marketplace owners do not clean the data and let their participants do quality checks. One interviewee explained that their data buyers rate the quality of the datasets that are sold at their data marketplace: *"We designed quality standards. With those standards, data is rated"* (DM2). Data buyers can refer to the quality ratings to estimate the data quality they can expect from a data set.

4.2.2.4 Data Storage

Data marketplace owners need to store data at a secure location that is scalable. Storage facilities have to cope with the volume, velocity and variety of data. As explained in section 4.2.1, data can be stored using a centralized or decentralized platform infrastructure. According to Curry (2016), distributed platform infrastructures are more suitable to scale data storage. Furthermore, decentralized data storage is perceived as more secure, because there is not one single point of control over the data (Ølnes et al., 2017). The data is distributed over several nodes, which avoids single points of failure.

4.2.2.5 Data Analysis

During data analysis, data marketplace owners can aggregate and analyze the datasets to extract new insights. Martens & Mueller-langer (2018) state that there is a need for aggregation of data across car brands in the automotive industry. One interviewee explained that they aggregate and analyze the data of their data sellers: *"we have many data suppliers. We process data, remove mistakes from the data, link data together and sell this as an aggregated product"* (DM7). The data analysis performed by this data marketplace owner is a task that would otherwise be performed by external developers. Some data marketplace owners do not analyze the data, because their customers analyze the data themselves. As one interviewee explained: *"we do not analyze the collected data to gain insights. That is something that our customers do"* (DM6). Another interviewee explained that they provide tools for their customers who use the tools for data analysis: *"Business users get visualization tools to derive the data into a graph tool. They can analyze the data and get the visuals"* (DM2). Deichmann et al. (2016) distinguish between data marketplace owners that are more technical and data marketplace owners that are more service focused. The data marketplace owners with a technical focus do not perform any data analysis and only forward data on their platform, whereas data marketplace owners who are service focused analyze data bundles to create insights.

4.2.2.6 Data Distribution

Similar to data collection, data distribution is an activity that all data marketplace owners perform. One data marketplace owner explained that data collection and distribution are their main activities: *"our data marketplace has two main functionalities. One is to show the meta-data of available datasets. The other is the brokerage functionality. That is to get data from a data provider and distribute this to all data users who need to subscribe to a data publication. So it's a data delivery and brokerage service"* (DM3). As explained in section 4.2.2.1, data sellers may restrict data usage for certain user groups or in certain areas. Depending on the contract conditions that were agreed upon during data collection data is sold to buyers.

4.2.3 Platform Infrastructure and Data Processing Activities as Business Model Dimensions, Part of Value Delivery

Based on the categories platform infrastructure and data processing activities, we derive identically named dimensions. We consider the *platform infrastructure* as a *resource* of data marketplace owners to *deliver value* to customers. Data marketplace owners store data in a centralized or decentralized platform infrastructure. In *centralized* platforms, data control shifts towards the data marketplace owner who manages the storage location. The data marketplace owner has access to the data that is traded at the data marketplace and can perform data analysis to transform the data. In *decentralized* platform infrastructures, the data seller maintains data control. Decentralized platform infrastructures do not have a central access point, which limits the data processing activities that a data marketplace owner can perform.

We identified six *data processing activities* in section 4.2.2. Some data marketplace owners explained that they perform *all* data processing activities to deliver value to their customers. They completely transform the data collected from their data sellers into a new data product to deliver to data buyers. Other data marketplace owners perform a *limited* number of processing activities. They focus on data collection and delivery to facilitate direct data transfer between data sellers and data buyers. Their customers process the data themselves. Depending on the activities that the data marketplace owner performs, their value proposition can be more service oriented, when all activities are performed, or technical oriented, when a limited number of processing activities are performed. We consider the dimension *data processing activities* as key activities for data marketplace owners, part of the meta-characteristic *value delivery*.

4.3 Revenue Model Leading to the Dimensions Revenue Streams and Data Pricing Mechanism

Data marketplace owners combine revenue streams and data pricing mechanisms to monetize data. We explain the revenue streams applied by data marketplace owners in section 4.3.1. Numerous interviewees indicated that it is challenging for the data marketplace owner and their customers to price data. In section 4.3.2 we explain how the price of the traded data is determined at data marketplaces. In section 4.3.3 we conclude by deriving revenue streams and data pricing mechanisms as business model dimensions that data marketplace owners apply to capture value.

4.3.1 The Revenue Streams to Generate Income

Data marketplace owners generate income from their revenue streams. A data marketplace owner can receive numerous revenue streams. For example, they may charge customers for the usage of their marketplace and customers can be charged for the data that is transferred. Muschalle et al. (2012) identify five revenue streams for data marketplaces:

- Free – the data marketplace owner provides the platform for free. Governments normally apply free revenue streams, because they are funded by tax money. Companies who want to attract more customers sometimes offer products or services for free. Muschalle et al. (2012) explain that when a substantial number of data buyers is active on the data marketplace, commercial data sellers are attracted to the data marketplace.

- Usage-based prices – the customers pay for a product or service per usage unit. This can be time based or volume based. Muschalle et al. (2012) explain that some data marketplace owners provide a consultancy service. They advise customers about data purchase or sale. This service is charged by a data marketplace owner per hour.
- Package pricing – data marketplace owners sell a combination of data goods or services for a fixed price. This has the advantage that the data marketplace owner can capture value from items that are worth little when they are sold individually.
- Flat fee tariff – there is one fixed fee to pay for the product. The data marketplace owner defines a period of time during which the customer may use the product or service. For example, they permit platform usage for an unlimited amount of time or restrict platform usage to a month or year for a fee.
- Freemium –the customers may use basic services for free and have to pay for services with additional value. Data marketplace owners often combine this revenue stream with another revenue stream. A customer can for example get a limited amount of data for free and has to pay when the limit is exceeded. One interviewee explained that they combine the freemium model with a usage based model (DM1). Developers can get data up to a limit of 250.000 transactions per month for free. When they exceed this limit, customers pay a price per data volume that is transferred.

4.3.2 Monetization of Data with Fixed or Dynamic Pricing Mechanisms

The data pricing mechanism specifies how the prices of the data that is traded are established. One interviewee explained that monetization of data is challenging: *"people do not know how to value data. This is a problem. You cannot have a marketplace where you do not know the value of what you are selling"* (DM4). The lack of intellectual property rights makes the value attribution of data difficult (Koutroumpis et al., 2017). Overall, data marketplace owners apply two types of pricing mechanisms. These are fixed pricing and dynamic pricing.

When fixed pricing mechanisms are applied, the data price is predefined and static. One interviewee explained that they trade data based on fixed prices: *"the data price is predefined and the total price is determined based on how much data the data seller consumed"* (DM6). Fixed pricing mechanisms are often applied in monopolies. Muschalle et al. (2012) explain that monopolies set a price in the data market and adjust the price depending on the buyers' willingness to pay for the data. Different prices for different customer segments can be applied, known as price discrimination (Muschalle et al., 2012). In the automotive industry, the data pricing mechanism is often set by the data seller (Martens & Mueller-langer, 2018). Martens & Mueller-langer (2018) explain that OEMs can fix a price for their data, because they have monopoly power. Third party service providers who are dependent on data of OEMs have to settle for the fixed price.

Dynamic pricing mechanisms are negotiated, auctioned or based on real-time market conditions (Muschalle et al., 2012; Osterwalder & Pigneur, 2010; Täuscher & Laudien, 2017). Data marketplace owners who apply dynamic pricing mechanisms aim for data sellers to become price takers. In order for dynamic pricing to succeed, an interviewee explained: "*There is a need for price discoveries and mechanisms that calculate liquidity based on the market and come up with the price. This is still very abstract"* (DM4). Although dynamic pricing is what this data marketplace owner strives for, they apply fixed pricing models in practice: *"Fixed pricing is the easiest play in the book. Come up with a number, and see if people are interested or drop the price. But this is definitely not the solution, because it's not in people's normal workflow to go and put a price on data. Nobody knows how to do this"* (DM4).

4.3.3 Revenue Streams and Data Pricing Mechanisms as Business Model Dimensions, Part of Value Capture

We derive two dimensions from the category 'revenue models'. The dimension *revenue streams* shows how the data marketplace owner generates income. Data marketplace owners can apply a combination of the following revenue stream models; free, usage based, package pricing, flat fee tariff or freemium model. The *data pricing mechanism* indicates how the prices of the data are established at data marketplaces. Fixed pricing mechanisms are static and set by the data seller, buyer or data marketplace owner. This is the pricing mechanism that is mainly applied by data marketplace owners in practice. Prices that are based on auctions, negotiations or real-time markets are examples of flexible pricing mechanisms. Such pricing mechanisms are desirable to establish competitive pricing at data marketplace with a market orientation. We include the dimensions *revenue streams* and *data pricing mechanisms* in our preliminary taxonomy, part of the meta-characteristic *value capture*.

4.4 Aligning Induced and Deduced Business Model Dimensions

Five business model dimensions are induced in section 4.1 – 4.3 that we include in our taxonomy to distinguish business models of data marketplaces: (i) contract, (ii) platform infrastructure, (iii) data processing activities, (iv) data pricing mechanism and (v) revenue streams. We supplement these dimensions with dimensions from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). Spiekermann (2019) and Fruhwirth et al. (2020) include dimensions that correspond to our induced dimensions, as shown in Figure 16. Furthermore, they define business model dimensions that we did not consider up to this point. We align these dimensions in the following paragraphs to compose our preliminary taxonomy.

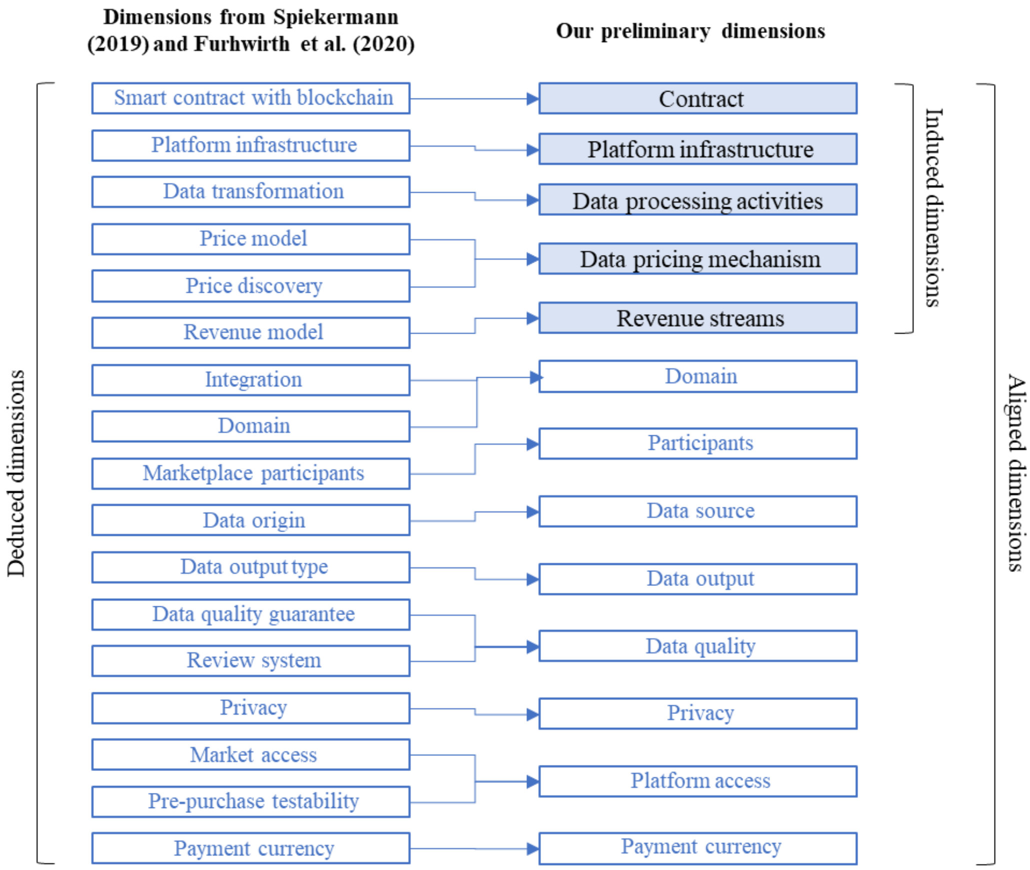

Figure 16: Business model dimension alignment

The dimensions *smart contract*, *platform infrastructure*, *data transformation*, *price model*, *price discovery and revenue model* from Spiekermann (2019) and Fruhwirth et al. (2020) correspond to the dimensions that we derived in section 4.1 – 4.3. The dimension *smart contract* is defined by Fruhwirth et al. Their dimension states whether data marketplaces have smart contracts or not. We include the preliminary dimension **contract** in which standardized and negotiated contracts are considered. It is significant that the dimension contract is not considered in the taxonomy of Spiekermann, whereas this is a dimension that appears to be key in forming customer relationships (see section 4.1). The *platform infrastructure* that Spiekermann and Fruhwirth et al. define in their taxonomy is identical to our preliminary dimension **platform infrastructure**. The dimension *data transformation,* defined by Spiekermann, corresponds to our preliminary dimension **data processing activities**. Spiekermann defines raw data trade, data standardization and data aggregation as the key transformation activities. However, as explained in chapter 4.2.2, we consider additional data processing activities. The dimension *price model* defined by Spiekermann and *price discovery* defined by Fruhwirth et al. correspond to our preliminary dimension **data pricing mechanism**. Likewise, the *revenue model* that Spiekermann defines corresponds to the preliminary dimension **revenue streams**.

Furthermore, Spiekermann (2019) and Fruhwirth et al. (2019) define dimensions that we did not consider thus far. These are the dimensions *integration, domain, marketplace participants, data origin, data output type, data quality guarantee, review system, privacy, market access, pre-purchase testability* and *payment currency*. We align these dimensions to form preliminary dimensions for our taxonomy. Spiekermann defines the *integration* of a data marketplace as domain-specific or domain-unspecific. A data marketplace with a specific domain has data from one industry. A data marketplace with an unspecific domain trades in data across multiple industries. The dimension integration overlaps with the dimension *domain* from Fruhwirth et al. They define domain as the information that the data assets contain. This can be finance, geo, address, sensor and personal. We merge both these dimensions into the preliminary dimension **domain** in our taxonomy. The domain concerns the market in which the data marketplace is active. The overarching domain in our research is the automotive market. More specific characteristics of this domain will be evident from the data marketplaces that we explore in chapter 5. The *marketplace participants*, as defined by Fruhwirth et al., refers to the data sellers and buyers who are matched at the data marketplace. This dimension will be included as preliminary dimension **participants** in our taxonomy. Similar to the dimension *domain*, more specific characteristics of the dimension *participants* will be explored in chapter 5. We alter the dimension *data origin* from Fruhwirth et al. into **data source**. The characteristics government, social media and commercial data sources are added to the characteristics from Fruhwirth et al. As described in appendix B, self-generated data is produced by the data marketplace owner. Data marketplace owners derive community data from other data marketplaces or crowdsourcing services. These are potential data sources for data marketplace owners. Furthermore, Fruhwirth et al. characterize the *data output* with different format types. As explained in section 4.2.2, the data output differs depending on the data processing activities that data marketplace owners perform. Data can remain unchanged and be transferred directly between the data seller and buyer or data can be processed by the data marketplace and sold as a transformed data product. Therefore, we adapt the characteristics of **data output** to aggregated data and standardized data. Next, we merge the dimensions *data quality guarantee* and *review system* from Fruhwirth et al. into one dimension. Fruhwirth et al. characterize the dimension data quality guarantee as *yes* or *no*. This does not add significant information to the data quality that a data marketplace owner ensures. Fruhwirth et al. define who evaluates the data quality in the dimension *review system*. We alter these dimensions into the preliminary dimension **data quality** with the characteristics *reviews by the marketplace* owner and *user reviews*. Furthermore, Fruhwirth et al. define the dimension *privacy* as part of the value proposition of a data marketplace. As explained in section 4.1.1 data privacy is regulated by the data marketplace owner. Thus, we include **privacy** as a preliminary dimension in our taxonomy. Moreover, Spiekermann (2019) defines the dimension *market access* as closed or open. In a closed market, a limited number of participants are allowed and in an open market the number of participants is broad and unknown. We name this dimension **platform access** in our preliminary taxonomy. Finally, the dimension **payment currency** from Fruhwirth et al. is considered in our preliminary business model taxonomy. Data marketplaces differ in terms of payments that are transferred in fiat currency or cryptocurrency.

We do not include all dimensions from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) in our taxonomy. The dimensions *market positioning*, *time relevancy* and *access type* are excluded. The *market positioning*, which Spiekermann defines, shows whether the platform is owned by an independent party or a buyer or seller. These characteristics are part of our definition of data marketplaces as explained in section 1.2.1. In our definition, data marketplaces have a private, consortium or independent ownership. Market positioning is not a dimension in our taxonomy. Fruhwirth et al. include the dimension *time relevancy* in their taxonomy. This dimension entails whether uploaded data is static or dynamic. If the time relevancy is dynamic, the data is updated regularly. This is a technical property that should be discussed as part of the data output dimension in the taxonomy. We do not consider this dimension as a stand-alone business model dimension. Finally, Fruhwirth et al. define the *access type* as part of value delivery of a data marketplace. They characterize access type with API, download or specialized storage. These characteristics are defined on a granular level. We do not demand that level of specificity to distinguish the business models of data marketplaces.

Table 16: Description preliminary dimensions

Dimension	**Description**
Domain	the market in which the data marketplace is active
Participants	the data sellers and buyers who are matched at the data marketplace
Data source	the governmental, social media, self-generated or community source where the data on the data marketplace is collected from
Data output	the aggregated or standardized data offering
Data quality	the user reviews or reviews by the data marketplace owner to guarantee data quality of the traded data
Privacy	the anonymization or encryption of data to protect data privacy
Contract	the negotiated or standardized agreements that regulate data trade
Platform access	the open or closed platform access for the customers
Platform infrastructure	the centralized or decentralized storage location of data
Data processing activities	the performance of all or a limited amount of activities by the data marketplace owner to increase the value of the traded data
Revenue streams	the manner in which the data marketplace owner generates income by applying usage based, package pricing, flat fee tariff or freemium models
Data pricing mechanism	the fixed (set by data marketplace owner, sellers or buyers) or dynamic (auction, negotiation, real-time market) pricing mechanism of the data output
Payment currency	the fiat or cryptocurrency in which payments are transferred

The dimensions that are discussed in this section and described in Table 16 constitute our preliminary taxonomy. Spiekermann (2019) and Fruhwirth et al. (2020) include varying dimensions in their taxonomies and do not provide an explanation why they include or exclude certain dimensions. We bridged this research gap by aligning their dimensions and explained why certain dimensions are excluded from our taxonomy. In addition, we based some of our dimensions on interviews with data marketplace owners. This adds credibility to our taxonomy and sets our taxonomy apart from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). Our preliminary taxonomy in Table 17 is still conceptual, which means that existing data marketplace cases are not yet classified in our taxonomy. In chapter 5 we refine our taxonomy with dimensions and characteristics from the business models of existing data marketplaces in the B2B automotive industry. Therefore, the dimensions and characteristics in our preliminary taxonomy are still subject to change.

Table 17: Preliminary taxonomy

	Component	Dimension	Preliminary characteristics					
Value creation	Customer segment	Domain	To be refined					
		Participants	To be refined					
	Value proposition	Privacy	Anonymized	Encrypted				
		Data source	Government	Social media	Commercial	Self-generated	Community	
		Data output	Transformed data	Non-transformed data				
		Data quality	Reviews by marketplace owner	User reviews				
	Customer relationship	Contract	Negotiated	Standardized				
Value delivery	Channels	Platform access	Closed	Open				
	Key resources	Platform infrastructure	Centralized	Decentralized				
	Key activities	Data processing activities	All	Limited				
Value capture	Revenue	Revenue streams	Free	Usage based	Package pricing	Flat fee tariff	freemium	
	Pricing model	Data pricing mechanism	Set by data marketplace owner	Set by sellers	Set by buyers	Auction	Negotiation	Real-time market
		Payment currency	Crypto	Fiat				

4.5 Conclusion of Chapter 4

In this chapter we explored business model categories to induce business model dimensions that distinguish the business models of data marketplace owners in practice. We aligned the induced dimensions with deduced dimensions from the business model taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) to create our preliminary taxonomy. This leads to an answer to the first sub-question *Based on an exploration of the business models of data marketplace owners in the B2B automotive industry and based on the existing generic business model taxonomies for data marketplaces, what dimensions can be derived to include in our preliminary business model taxonomy?* By applying the grounded theory method, we identified five categories that were main topics of discussion during our interviews with data marketplace owners. The categories are data regulation, customer relationships, platform infrastructure, data processing activities and revenue models. From these categories, five dimensions are induced that need to be included in our taxonomy to distinguish the business models of data marketplaces. We induced the dimension *contract* from the categories data regulation and customer relationships. Data marketplace owners apply negotiated or standardized contracts to incorporate data regulation into their business model and create customer relationships. The dimensions *platform infrastructure* and *data processing activities* are induced from the identical categories. The centralized or decentralized platform infrastructure are a resource for data marketplace owners to perform all or limited data processing activities. This impacts the value proposition that is delivered to the customer. Last, we induced the dimensions *revenue streams* and *data pricing mechanism* from the category revenue models. Pricing of data is challenging for data marketplace owners and their customers, because people are not used to monetize data. Owners of data marketplaces with a hierarchical orientation often apply fixed pricing mechanisms and owners of data marketplaces with a market orientation aim for dynamic pricing mechanisms, but often opt for fixed pricing. Finally, we deduced business model dimensions from the existing business model taxonomies for data marketplaces and aligned all dimensions in section 4.4. This results in our preliminary taxonomy. In the next chapter, our preliminary taxonomy is further refined with business model characteristics that we induced from the data marketplace cases in the B2B automotive industry.

5. Taxonomy Classification

In this chapter the characteristics of the business models of TomTom (TT), INRIX (IN), HERE (HE), Caruso (CR), IOTA and Ocean Protocol (OP) are described. We induced the business model characteristics from the case websites, terms and conditions, whitepapers and external sources of the data marketplaces by applying content analysis as explained in section 3.3. According to Nickerson et al. (2013), each dimension in a taxonomy must contain characteristics that are mutually exclusive and collectively exhaustive. The mutually exclusive rule means that "no object can have two different characteristics in a dimension" (Nickerson et al., 2013, p. 341). We classified data marketplaces under the assumption that each data marketplace has one business model. Thus, each data marketplace is classified in one business model characteristic per dimension.

Our refined taxonomy in Table 18 shows the business model characteristics that owners of data marketplaces in the B2B automotive industry apply. With respect to our preliminary taxonomy, we made the following structural changes. First, the dimension *data service* is added. The data service that data marketplace owners perform shapes their value proposition and is an integral part in their business model. Second, the dimension *data source* is removed because the characteristics that we identified for this dimension overlapped with the dimension *participants*. Third, we added newly induced characteristics from the case documents to our taxonomy and removed characteristics that are not represented in the business models of the data marketplace cases. For example, none of the data marketplaces have real-time pricing mechanisms or data auctions. These are "null characteristics" and are removed in our refined taxonomy, in line with OE3 (see Table 1).

Table 18: Refined taxonomy

	Component	Dimension	Characteristics				
Value creation	Customer segment	Domain	Location (TT, IN, HE)	Automotive (CR)	All industries (IOTA, OP)		
		Participants	Data sellers, data buyers, internal & external developers (TT, IN, HE)	Data sellers, data buyers & external developers (CR, IOTA, OP)			
	Value proposition	Data service	Customized map service (TT, IN)	Data brokering service (CR, IOTA, OP)	Both (HE)		
		Data output	Aggregated data (TT, IN)	Standardized data (CR, IOTA, OP)	Both (HE)		
		Data quality	Reviews by marketplace owner (TT, IN)	User reviews (IOTA, OP)	No info (HE, CR)		
		Privacy	Anonymized (TT, IN, CR)	Encrypted (HE, IOTA, OP)			
	Customer relationship	Contract	Negotiated (TT, IN, CR)	Standardized (IOTA, OP)	Both (HE)		
Value delivery	Key channels	Platform access	Closed (TT, IN, CR)	Open (HE, IOTA, OP)			
	Key resources	Platform infrastructure	Centralized (TT, IN, HE, CR)	Decentralized (IOTA, OP)			
	Key activities	Data processing activities	All (TT, IN, HE)	Limited (CR, IOTA, OP)			
Value capture	Revenue	Revenue streams	Usage based (TT, IN)	Usage based & freemium (HE)	Commission (CR)	Donations (IOTA)	No info (OP)
	Pricing model	Data pricing mechanism	Set by marketplace owner (TT, IN)	Set by data seller (CR, IOTA, OP)	Both (HE)		
		Payment currency	Fiat currency (TT, IN, HE, CR)	Cryptocurrency (IOTA, OP)			

In the following sections 5.1 to 5.3, we describe the characteristics within each dimension separately. Together they make up the meta-characteristics value creation, delivery and capture. In section 5.4 we combine the information from the previous sections to answer the second sub-question *What business model characteristics do owners of current data marketplaces in the B2B automotive industry apply?*

5.1 Value Creation

In the following paragraphs we describe the identified characteristics of the business model dimensions domain, participants, data service, data output, data quality, privacy and contract. These dimensions are part of the meta-characteristic value creation.

5.1.1 Domain

A data marketplace owner targets a *domain*. The domain entails the market in which the data marketplace is active. The characteristics of a data marketplace domain are *location*, *automotive* and *all industries*.

TomTom, INRIX and HERE focus on one specific domain in the automotive industry; location. Location technology vastly changed the design of today's maps. People used to rely on static maps to navigate themselves. By use of location technology, TomTom, INRIX and HERE design dynamic maps and communicate real-time road conditions to their customers. They envision the realization of an autonomous and connected world through location technology. Autonomous driving is a development which requires location technologies: "Mobility will be autonomous, connected and electric. TomTom's leading location technologies are accelerating this revolution, together with drivers, cities and our partners" (TT-4). HERE responds to trends in autonomous driving and IoT: "From autonomous driving, to the Internet of Things, we are building the future of location technology through strategic partnerships" (HE-1a). These data marketplaces are specialists in the domain location, a domain that is part of the automotive industry.

Caruso focuses on the complete automotive domain. This domain includes numerous segments such as the (i) vehicle position, movement and surroundings (ii) vehicle health and maintenance (iii) vehicle non-powertrain hardware (iv) vehicle powertrain resources (v) vehicle powertrain hardware (vi) mobility services and (vii) auxiliary devices (CR-3). These segments appeal to a variety of stakeholders from markets that are connected to the automotive industry such as design, development, manufacturing, marketing and sales of vehicles. With a focus on the automotive domain, Caruso aims to support the digital transformation in the entire automotive industry.

IOTA and Ocean Protocol target a variety of industries. They do not limit themselves to only the automotive industry. IOTA targets the supply chain, smart cities, energy, manufacturing and healthcare industry. Ocean Protocol reasons that a cross-industry focus is required to advance use cases in a specific industry. For example, the development of autonomous vehicles (AVs) relies on artificial intelligence (AI), which runs on algorithms. Algorithms need to be trained on high volumes of data to be improved: "AV training data illustrates how not all data is fungible: a mile driven in a blizzard is worth more than a mile driven on an empty, sunny desert highway" (OP-3a). Training data concerning weather conditions and transportation infrastructures come from multiple industries. By trading data across industries, these data marketplace owners aim to create the most disruptive impact.

5.1.2 Participants

The *participants* are the actors who are matched at a data marketplace to trade data. The characteristics that we identified of the data marketplace participants are *data sellers*, *data buyers*, the *marketplace owner*, *internal developers* and *external developers*.

There is little distinction in the specific data sellers, data buyers and external developers who are targeted by the data marketplaces. Large companies such as OEMs, Tier 1 suppliers and mobile broadband companies with much data, are qualified to be both data sellers and data buyers. All data marketplaces name corporates like BMW, Bosch and Vodafone as their participants. Smaller enterprises, such as start-ups or individual developers that aim to develop services with the data provided on the platform, are also both data sellers and buyers at the data marketplaces. The data sellers and data buyers in general are similar among the targeted companies and no distinction is made between the two in the characteristics of the participants. There is however a variation in the type of developers involved in the business models, leading to the categorization of participants in internal and external developers compared to only external developers.

TomTom, INRIX and HERE have internal developers. They use the data traded at their data marketplace for their own development purposes. With proprietary algorithms the inhouse analysts create the value proposition these data marketplace owners offer: "Using powerful analytical tools, TomTom can now precisely tailor offerings to deeply articulated customer segments" (TT-4a). Thus, the roles 'data marketplace owner' and 'third party service provider' are merged. In addition to internal developers, external developers participate on these platforms. The external developers make use of the marketplace datasets and develop applications that are integrated in the offering of the data marketplace: "Automakers can employ OpenCar's cloud environment which crafts a foundation for app content developers to seamlessly build and deploy newer content" (IN-3). The internal developers develop the value proposition of the data marketplace and the external developers use the data marketplace to provide their own service at the data marketplace.

Caruso, IOTA and Ocean Protocol do not have internal developers who transform data, but focus on external developers. IOTA targets external developers who have a specific need for IoT data and Ocean Protocol targets AI developers who need data to test and improve their algorithms. Ocean Protocol identifies a gap between the knowledge of start-ups and larger enterprises: "AI startups have amazing algorithms but are starving for data; and typical enterprises are drowning in data but have less AI expertise" (OP-3a). The participants who IOTA and Ocean Protocol target are external developers who further process the datasets themselves. On these data marketplaces, the roles of marketplace owner and third party service provider are separated.

5.1.3 Data Service

The *data service* specifies what services the data marketplace owner offers to the customer to create value. The data service characteristics are *customized map service*, *data brokering service* and *both*.

TomTom and INRIX provide a customized map service to trade aggregated data. They aggregate the data from their participants into mapped data. The customized map service is subdivided in (i) real-time traffic service (ii) electrical vehicle (EV) service (iii) parking service and (iv) speed camera service, to compose maps (TT-1, IN-1). Through a real-time traffic service, the data marketplace owner provides information about congestion, arrival time and routes to a location. In the EV service the data marketplace owner optimizes routes for EVs. This includes information about the nearest charging station and the occupation of charging stations. The parking service shows the availability of parking spots and the speed camera service alerts drivers about speed controls. The customers of TomTom and INRIX can choose the service that fulfills their needs. As such, customers receive different services. For example, Renault and BMW are both customers of TomTom (TT-1). However, their in-car navigations differ from one another. Renault makes use of the real-time traffic and speed camera services from TomTom. TomTom provides BMW with EV navigation in their cars. The location services from TomTom can be acquired separately and are tailored to the needs of the customer. BMW is also a customer of INRIX for its on-street parking service (IN-1). The on-street parking service provides BMW cars with information about free parking spots. Data marketplace owners who perform the customized map service provide external developers with software development kits (SDK) and access to Application Programming Interfaces (APIs). Developers can use the SDKs and APIs to develop their own web applications. The services or apps that external developers create are integrated with the customized map service of the data marketplace owner to extend their service (TT-2a). The data marketplace owners train their participants to use their tools. This requires an investment in time and money from the data marketplace owner (IN-5). However, the data marketplace owner benefits from this investment because it raises the entrance barriers for competitors and creates lock-in effects.

Caruso, IOTA and Ocean Protocol provide a data brokering service to enable data trade between their participants. This service includes minimal interference of the data marketplace owner: “We neither offer data nor services ourselves. We focus on providing the best possible brokering infrastructure” (CR-1). They do not change the content of the data from their participants, but standardize the data into one format. The data marketplace provides the technical infrastructure for direct trade between the data seller and buyer. The data marketplace owner educates their participants about the use of the technical infrastructure and about possible use cases with the data from the data marketplace through online documentation. In addition, the data marketplaces provide contracts to ensure secure data trade. Template licenses are used by the participants to trade their own datasets. This is further specified in section 5.1.7.

HERE offers both the customized map service and data brokering service. These services comprehend two different value propositions. Their customized map service is similar to the service that TomTom and INRIX provide. In addition, HERE provides the data brokering service to exchange data directly between the data seller and buyer: “In the HERE Open Location Platform, data consists of both maps and location information that HERE provides, such as Real-Time Traffic and Weather, as well as data that you and other users provide” (HE-1b). Both services generate different data outputs; aggregated and standardized.

5.1.4 Data Output

The *data output* of a data marketplace shows what data the data marketplace owner trades. Its characteristics are *aggregated data*, *standardized data* and *both*.

TomTom and INRIX trade in aggregated data. This is related to the data service that they perform. The customized map service generates mapped data which is aggregated. The created data consists of up-to-the-minute information on location, speed and directional heading of vehicles. As stated by TomTom: “the data is an essential ingredient of TomTom Personal Navigation Devices and online maps, and is branded as IQ Routes” (TT-3a). The maps and navigation software are aggregated products based on the data that their data sellers provide.

Caruso, IOTA and Ocean Protocol trade standardized data. Since these data marketplaces do not alter the content of their data sellers, the traded data is barely transformed. The data can be standardized by the data marketplace owner, as Caruso does, or the data can be standardized by the data seller participants. Caruso integrates the IT systems of their participants into their platform infrastructure to create a standardized format (CR-3). At Ocean Protocol and IOTA the participants standardize the data in schemes that are provided by the data marketplace owner. Metadata about the datasets is published at the data marketplace. The metadata contains information such as name, dateCreated, author, license, price and files are included in the metadata (OP-3a). Data buyers use this information to search for datasets at the data marketplace.

Similar to the explanation in section 5.1.3, HERE offers both aggregated and standardized data output. HERE produces aggregated data by performing their customized map service. The standardized data is the output of their data brokering service: “our platform provides comprehensive mapping content, an integrated suite of solutions, services and development tools and a marketplace for data to solve your complex location-based problems” (HE-1a).

5.1.5 Data Quality

Data quality entails who controls and preserves the quality from the data seller. The characteristics that we identify are *reviews by the marketplace*, *user reviews* and *no information*.

TomTom and INRIX review the data quality themselves. INRIX takes an active role in ensuring high quality datasets by combining data using proprietary algorithms: “INRIX combined this data with proprietary algorithms to produce much higher quality” (IN-5). When raw datasets from data sellers, such as GPS data, include noise they need to be separated and filtered. For example, TomTom describes the need to create speed patterns and separate road drivers from train travelers in GPS data: “a typical speed pattern appears when calls are coming from trains, because all handsets have the same speed and handover events, these data can be taken out” (TT-3a). Speeds of people on trains distort datasets that the data marketplace owner uses to analyze speeds of car drivers. The data marketplace owner filters and removes data that matches speed patterns of people on trains from the datasets.

IOTA and Ocean Protocol let their users review the data. IOTA makes use of their decentralized infrastructure to create a network of reviewers: “In order to make a transaction in the Tangle, two previous transactions must be validated with the reward for doing so being the validation of your own transaction by some subsequent transaction” (IOTA-1b). Ocean Protocol makes use of a similar system where a peer to peer consensus network is created between Ocean Keepers (OP-3b). Keepers are data marketplace users who are assigned the task to run nodes in the network. As a reward for chain keeping they receive Ocean Tokens (OP-3a). The data marketplace owners are not directly involved in preserving the data quality. The community reviews the data quality for a reward.

5.1.6 Privacy

Privacy indicates how stored data at a data marketplace is protected. Data marketplace owners implement privacy mechanisms that we characterize as *anonymized* and *encrypted*.

TomTom and INRIX anonymize the GPS data they collect. Caruso trades in anonymized data as well to adhere to privacy laws: “As a rule, the analyses of your browsing patterns are conducted anonymously; i.e. the browsing patterns cannot be traced back to you” (CR-2a). Through data anonymization, data marketplace owners safeguard individuals’ privacy.

The data that is stored at HERE, IOTA and Ocean Protocol is encrypted. Encrypted files are exchanged between data sellers and buyers to increase the data security. Only users with access to a decryption key are able to read the files: “The Secret Store checks the user’s authorization permissions on-chain to determine if they are authorized to decrypt a document. The Secret Store won’t allow decryption if the user doesn’t have authorization” (OP-3a).

5.1.7 Contract

The *contracts* that data marketplace owner manages define the agreement that enforces data trade between the data seller and data buyer. The contract characteristics that we identified are *negotiated*, *standardized* or *both*.

INRIX, TomTom and Caruso have negotiated contracts with their participants. The data marketplace owners negotiate contracts bilaterally, which results in separate agreements with each of their participants. The data sellers determine what data is delivered and for what purpose the data may be processed. Negotiated contracts with data sellers are usually long-term arrangements. For example, contracts with OEMs are on average three years long (IN-5). When the contract is signed and the data is delivered, the ownership over the sold data is transferred from the data seller to the data marketplace owner: “TomTom acquires ownership of the products the moment they are delivered in accordance with the contract or otherwise become available to TomTom” (TT-2b). The data marketplace owner translates the contract conditions into their terms of use to ensure the data is used as was agreed upon. Overall, negotiated contracts demand close partner communication, which requires high effort from the data marketplace owner. The data marketplace owner communicates with each data seller and buyer to understand their needs and offer personal assistance: “We'll dive in deep and figure out all the necessary details & synchronize on tasks, backlog and timelines” (CR-1). This enables the data marketplace owner to have deep understanding of their customers and their preferences.

Data marketplace owners with standardized contracts enable efficient data trade. Standardized contracts minimize the interaction with an intermediary party, which should decrease costs for the data marketplace owner and participants (IOTA-1b). IOTA and Ocean Protocol automate their contracts with the use of smart contracts. These are decentralized service agreements, stored in the DLT. Template service agreements can be altered by the data sellers and buyers to form a unique service agreement. The expected service and price are defined in the smart contract. Participants can solve disputes by reviewing the verifiable contracts. The trust in the intermediary is transferred to the ledger (IOTA-1b). By removing this third-party dependency, the data marketplace owners aim to create efficiency gains, innovation opportunities and stimulate the creation of new value propositions for their participants (IOTA-1b).

HERE offers both the negotiated and standardized contracts. Their subscriptions serve as contracts between the participants. HERE distinguishes between customized and commercial subscriptions (HE-1b). The customized subscriptions embody negotiated contracts. The data buyers are allowed to browse the marketplace for data, but they need to negotiate the data subscription offline with the data seller. Once a contract is negotiated, the data buyers are allowed access to the datasets. The commercial subscriptions serve as standardized contracts. Data sellers use subscription schemes to create a listing of their datasets with standardized pricing and usage terms. With the standardized contracts, data sellers and buyers can trade data without personal assistance of the data marketplace owner.

5.2 Value Delivery

In the following paragraphs we describe the characteristics that we identified for the dimensions platform access, platform infrastructure and data processing activities, as part of the meta-characteristic value delivery.

5.2.1 Platform Access

The *platform access* of a data marketplace concerns the degree of openness for participants to enter the platform. The characteristics of the platform access are *closed* and *open*.

TomTom, INRIX and Caruso have closed platform access. They restrict access to their platform with identity and access management. The users must authenticate themselves with company details and specifications about their data use. The platform owner approves or declines the registration requests.

HERE, IOTA and Ocean Protocol have open platform access and allow anyone to upload and buy data from the marketplace. Users can directly enter the data marketplace after they created a user account. It does not need to be approved by the data marketplace owner. As Ocean Protocol states: “The marketplaces built on Ocean Protocol will allow data to be accessed by all participants, ensuring that no central player can control or exploit the data” (OP-4a). The data marketplace owner does not restrict anyone from using their platform.

5.2.2 Platform Infrastructure

The *platform infrastructure* indicates how data is stored at the data marketplace. The platform infrastructure can be *centralized* or *decentralized*.

TomTom, INRIX, HERE and Caruso have a centralized platform infrastructure and store data on a central location. The central data storage facilitates up-to-date data and data standardization. INRIX explains: “a cloud environment creates the ability for app developers and trusted app brands to deliver up-to-the-minute contextual content due to its standardized functionality” (IN-3). The cloud, the central storage location of the data marketplace, is easily linked to clouds of other companies. HERE links their cloud to companies’ on-premise storage: “in the hybrid approach, an organization will use some services from the public cloud and deploy related services on its own servers, depending on the requirements of the location application and use case” (HE-3a). This offers enterprises the opportunity to store sensitive data on-premise and upload other data to the central cloud storage of HERE. In section 4.2 we explained that central storage benefits the data processing capabilities of the data marketplace owner. This is confirmed in the HERE terms and conditions that state: “submitting, posting or displaying Your Content in the HERE Services or otherwise providing Your Content to HERE, you grant HERE and its affiliates a perpetual, irrevocable, worldwide, royalty-free, non-exclusive, sub-licensable license to reproduce, adapt, modify, translate, publicly perform, publicly display, distribute, process and transfer your Content” (HE-2).

IOTA and Ocean Protocol have a decentralized platform infrastructure and store data across locations. There is no central administrator who controls the data. An example of the decentralized platform infrastructure is the blockchain technology that Ocean Protocol deploys. Each block in the blockchain functions as a point of control to secure the data (OP-4b). However, a fee must be paid per data transaction in blockchain. IOTA reasons that in the IoT industry the usage of blockchain is inefficient for micropayments: “paying a fee that is larger than the amount of value being transferred is not logical” (IOTA-3a). Therefore, IOTA designed their own DLT, called the Tangle, that has no transaction fees. In the Tangle, each participant has to approve two previous transactions to join the network. Overall, Ocean Protocol and IOTA aim to increase data transparency and data sovereignty by deploying a decentral infrastructure.

5.2.3 Data Processing Activities

In the *data processing activities* we define the activities performed by the data marketplace owner to increase the value of the data. The data processing activities can cover *all* activities of the data value chain or a *limited* number of activities.

TomTom, INRIX and HERE apply all data processing activities (see Figure 17). They start with intensive data collection to gather separate data points. Next, supporting systems, such as salesforce, are integrated in their architecture (TT-4c). The data marketplace owner standardizes the data transmitted from these systems in a data fusion engine (TT-4c). TomTom only wants to keep the necessary data of all collected datapoints. Therefore, they clean the data. This entails data filtering, separation and removal. The data marketplace owners store data in the cloud with a data update frequency of 3 minutes (TT-3a). Subsequently, they enrich data from road sensors with weather information to improve routing data. To close data gaps, sensor data is enhanced with GPS data to predict new datapoints (INRX-5). For data prediction, the data marketplace owner compares real-time traffic data to historic data (TT-3a). Routing data and estimated times of arrival (ETA) are distributed by the data marketplace owner to the customer. Data marketplace owners that perform all data processing activities make significant investments in their algorithms and data processing capabilities. INRIX, for example, invested $35 million in their real-time and predictive algorithms (IN-5). This increases the entry barrier for competitors.

<u>All</u>

1. Data collection
- Monitor user contributed data
- Collect own data
- Crowd-source vehicle data
- Collect GPS data
- Collect data over long time period

2. Data standardization
- Connect supplier software

3. Data cleansing
- Filter unnecessary data
- Separate data
- Remove data

4. Data storage
- Update content

5. Data analysis
- Enhance with road data
- Enhance with weather information
- Make algorithms
- Develop applications and services
- Apply proprietary algorithms
- Predict data
- Create speed profiles
- Estimate arrival time

6. Data distribution
- Provide mapped data

Figure 17: all processing activities (based on the induced codes in appendix E.7)

Caruso, IOTA and Ocean Protocol perform a limited number of data processing activities (see Figure 18). They do not clean the data and are limited in their data analysis. For example, Caruso's main activities are to collect, harmonize and distribute data (CR-3). During harmonization, they integrate different systems and standardize data. The data marketplace owners and their participants use schemas to structure data and ease data sharing between data traders. Structured metadata about the datasets is stored in a data catalog (CR-3, OP-3a, IOTA-1). The data that the data marketplace owners analyze are the platform usage patterns to decide about long term governance (OP-3b). Through the platform infrastructure, the data marketplace owners distribute the data directly between data sellers and buyers.

Limited

1. Data collection
2. Data standardization
   Integrate systems
   Provide data schemas
3. Data storage
   Store metadata in catalogs
4. Data analysis
   Analyze use patterns
   Report data metrics
5. Data distribution

Figure 18: Limited activities (based on the induced codes in appendix E.7)

5.3 Value Capture

A data marketplace owner captures value from the products and service that they sell. In the following paragraphs we explain the dimensions revenue streams, data pricing mechanism and payment currency within the meta-characteristic value capture.

5.3.1 Revenue streams

The *revenue streams* indicate how the data marketplace owner generates turnover. The characteristics of the revenue streams of data marketplaces are *usage based*, *usage based & freemium*, *commission*, *donations* and *no info*.

TomTom and INRIX receive usage based revenue streams. They create turnover by charging their participants for the usage of their data. TomTom charges their external developers per 1000 data transactions (TT-1b). INRIX charges OEMs based on their number of subscribers per month (IN-5).

HERE combines a usage based and freemium model. Participants can get up to 250.000 data transactions per month for free. When they exceed this limit, the data marketplace owner charges an add-on of $45 per month or a pro subscription of $449 per month to allow data access.

Caruso allows their participants to choose themselves what data pricing mechanism they prefer. Caruso receives a commission of the data that is sold at their marketplace (CR-3). The data marketplace owner defines the commission in the contract with the data seller.

IOTA is a non-profit organization and provides their platform for free. The organization is funded by donations from individuals and enterprises to maintain their platform (IOTA-1). IOTA also receives grants from governments to perform research.

Ocean Protocol owned by a non-profit organization as well but does not specify what their revenue streams are. Ocean Protocol allows their data sellers to choose any pricing strategy to sell their data, but Ocean Protocol does not receive revenue from the data that is sold.

5.3.2 Data pricing mechanism

The *data pricing mechanism* indicates how prices of the data are established between the trading entities. Prices can be *set by the marketplace owner*, *set by the sellers* or *both*. These pricing mechanisms are examples of fixed pricing mechanisms. We did not observe dynamic pricing mechanisms in the data marketplaces that we researched.

TomTom and INRIX sell their own data for which they set the price themselves. At Caruso, IOTA and Ocean Protocol the data sellers set the price for their data that they trade. HERE applies both pricing mechanisms. They set the price for their own aggregated data and their data sellers set the price for the standardized data that is traded.

5.3.3 Payment currency

The *payment currency* is the currency in which the payment is transferred. The characteristics are the *fiat currency* and *cryptocurrency*.

TomTom, INRIX, HERE and Caruso use fiat currency. When a data marketplace owner maintains fiat currencies, data can be traded in multiple currencies. In the Netherlands it is the Euro, in the United Kingdom the Pound is used and in the United States participants trade data using the Dollar as currency. Ocean Protocol and IOTA have their own cryptocurrency. These cryptocurrencies are called tokens, and can be used only at their data marketplace.

5.4 Conclusion of Chapter 5

Through describing the characteristics applied by the business models within each dimension, the second sub-question can be answered. *What business model characteristics do current data marketplaces in the B2B automotive industry apply?* The various data marketplaces researched in this paper; TomTom, INRIX, HERE, Caruso, IOTA and Ocean Protocol each apply a specific combination of characteristics to create, deliver and capture value.

TomTom and INRIX are both privately owned companies that trade data using a hierarchical orientation. Both companies apply location technology to improve navigation and mobility services focusing on one specific automotive domain; location. Their internal developers perform all data processing activities to create the value proposition. Through their customized map service, anonymized, aggregated data output is created, which is reviewed by the data marketplaces themselves. The agreements that enable trade between the participants are established through negotiated contracts. TomTom and INRIX deliver value using a central platform infrastructure with closed platform access. Finally, value is captured through usage based revenue streams where the prices are set by the marketplace owners and are paid in fiat currencies.

IOTA and Ocean Protocol apply opposite characteristics in their business model when compared to TomTom and INRIX. Both IOTA and Ocean Protocol can be described as independent data marketplaces that trade data in a market orientation. They target a wide variety of industries and do not limit themselves to the automotive industry. These data marketplaces do not have internal developers who process data but instead target external developers, who need data to improve their algorithms, to process the data themselves. The service these marketplaces provide can be described as a data brokering service, where data is encrypted and standardized in one format and exchanged directly between data sellers and buyers with minimal interference of the marketplace owner. Data quality is reviewed by the users themselves and no additional data processing activities are performed by the data marketplace owners. Efficient data trade is enabled using standardized contracts. An open, decentralized platform infrastructure allows anyone to upload and buy data from the marketplace. Data sellers set the prices themselves which are paid in cryptocurrencies. For IOTA, value is captured from donations and Ocean Protocol has not specified their revenue streams. They both do not receive direct revenue from the data that is sold.

HERE is a consortium owned data marketplace that applies a mixed hierarchical and market oriented trading structure. Their focus is set on the domain ‘location’. HERE creates two different types of value propositions. Both a data brokering service and a customized map service are offered to the customer. In the first, standardized data is traded directly between the data seller and buyer. In the second, aggregated data is created and improved by internal developers, which is eventually sold to the data buyer. Consequently, two types of contracts are enforced to enable data trade; negotiated and standardized contracts. The price of the aggregated data is set by the marketplace owner. In contrast, the price of the standardized data is set by the data seller themselves. Data is traded at an open, centralized platform. HERE combines the usage based and freemium model as their revenue model where prices are paid in fiat currency.

Caruso is also a consortium owned data marketplace with a mixed hierarchical and market orientation. Caruso facilitates data trade in the complete automotive domain. They do not process the data that is traded, but offer a data brokering service to allow direct exchange in standardized data between their participants. Through bilaterally negotiated contracts, their participants are assisted in data exchange. A closed, centralized platform infrastructure facilitates the data delivery. The data seller sets the price in fiat currency for the data that is sold. For every data sale, Caruso receives a commission.

In chapter 6 these cases will be further compared pairwise. TomTom and INRIX will be paired based on similar characteristics as well as IOTA and Ocean Protocol. HERE and Caruso are compared as dissimilar data marketplaces to eventually generate business model archetypes.

6. Taxonomy Patterns

Our taxonomy is designed to compare cases based on similarities and differences in their business models. In this chapter a pairwise comparison is performed, as explained in section 3.4, to generate business model archetypes. An archetype comprises business models that show similar reoccurring patterns. In chapter 6.1 we try to break simplistic frames by pairwise comparison. In Table 19 cases are paired based on the characteristics they share using color-coding. The blue color represents similar characteristics of TomTom and INRIX and the green color represents similar characteristics of IOTA and Ocean Protocol. HERE and Caruso only share two identical characteristics, indicated with bold lettering. In section 6.1 comparisons are made between TomTom and INRIX and between IOTA and Ocean Protocol because of their corresponding characteristics. In these cases focus is put on the dissimilarities between these cases to break the simplistic frames. In contrast, HERE and Caruso, who only share two characteristics and are therefore seemingly dissimilar cases, are compared based on their similarities. These comparisons show whether the paired cases should or should not be grouped together in a business model archetype. Our business model archetypes are presented in section 6.2.

Table 19: Similar taxonomy characteristics

	Component	Dimension	Characteristic				
Value creation	Customer segment	Domain	Location (TT, IN, HE)	Automotive (CR)	All industries (IOTA, OP)		
		Participants	Data sellers, data buyers, internal & external developers (TT, IN, HE)	Data sellers, data buyers & external developers (CR, IOTA, OP)			
	Value proposition	Data service	Customized map service (TT, IN)	Data brokering service (CR, IOTA, OP)	Both (HE)		
		Data output	Aggregated data (TT, IN)	Standardized data (CR, IOTA, OP)	Both (HE)		
		Data quality	Reviews by marketplace owner (TT, IN)	User reviews (IOTA, OP)	No info (**CR, HE**)		
		Privacy	Anonymized (TT, IN, CR)	Encrypted (HE, IOTA, OP)			
	Customer relationship	Contract	Negotiated (TT, IN, CR)	Standardized (IOTA, OP)	Both (HE)		
Value delivery	Key channels	Platform access	Closed (TT, IN, CR)	Open (IOTA, OP)	Both (HE)		
	Key resources	Platform infrastructure	Centralized (TT, IN, **HE, CR**)	Decentralized (IOTA, OP)			
	Key activities	Data processing activities	All (TT, IN, HE)	Limited (CR, IOTA, OP)			
Value capture	Revenue	Revenue streams	Usage based (TT, IN)	Usage based & freemium (HE)	Commission (CR)	Donations (IOTA)	No info (OP)
	Pricing model	Data pricing mechanism	Set by data marketplace (TT, IN)	Set by seller (CR, IOTA, OP)	Both (HE)		
		Payment currency	Fiat currency (TT, IN, HE, CR)	Cryptocurrency (IOTA, OP)			

Seemingly similar characteristics of TomTom and INRIX

Seeming similar characteristics of IOTA and Ocean Protocol

6.1 Breaking Simplistic Frames

In this section the pairwise comparison is performed. In section 6.1.1 we search for dissimilarities between TomTom and INRIX, in section 6.1.2 we search for dissimilarities between IOTA and Ocean Protocol and in section 6.1.3 we search for similarities between HERE and Caruso. After each pairwise comparison, an explanation is provided whether the cases should be grouped in one business model archetype based on similarity or should be separated based on dissimilarity.

6.1.1 Dissimilarities between TomTom and INRIX

TomTom and INRIX are equal in their goal to create intelligent maps by use of location analytics. This shows in their similar taxonomy characteristics such as the focus on the location domain, the performance of a customized map making service and usage of a centralized platform infrastructure to store and analyze their data. However, dissimilarities are present between the two and our analysis leads to interesting insights in business model archetypes.

First, in the value creation TomTom and INRIX differs in the *participants*. TomTom puts emphasis on self-generated data whereas INRIX relies on commercial data sellers. The self-generated data of TomTom originates from their branded smart devices: "the core sources of traffic data collection systems are the probe data from the cell phone operators in the various countries as well as GPS probes from the installed base TomTom connected devices and commercial fleets" (TT-3). INRIX has put its focus on location and connected services which they employ to generate a crowd-sourcing network: "we developed the idea of partnering with commercial fleets, that were already in the process of installing GPS devices in their vehicles" (IN-5). They had over 200 B2B customers in 2012, which put them ahead of TomTom who had less than 20 B2B customers at the time (IN-5). The data of INRIX mainly comes from their corporate data sellers. Thus, TomTom has more emphasis on data from their own hardware devices and INRIX relies on the data from their corporate participants.

Second, INRIX and TomTom differ in their implementation of the *centralized platform infrastructure* and performed *data processing activities* to deliver value. TomTom and INRIX basically adhere to the same principle of centralized data storage but there is a difference in their partnerships with cloud providers. TomTom and INRIX both form partnerships with Microsoft to store their data in the Azure cloud, but INRIX also partners with Amazon and makes use of the AWS Cloud. INRIX integrates the AWS Cloud into their OpenCar service to connect with Amazon's conversational technology, called Alexa (IN-3). Through this service, drivers are able to instruct Alexa for executing tasks instead of using their in-vehicle dashboards. This product integration is envisioned for more luxury cars (IN-3) and appeals to the high-end OEMs. Furthermore, the algorithms of TomTom and INRIX are proprietary. They produce their data output by applying these algorithms. This influences the accuracy and speed of their data delivery. Although TomTom and INRIX are similar in their data processing activities, they distinguish their activities with their algorithms and compete for best delivery of location information.

Last, there is a difference in the *revenue streams* TomTom and INRIX have to capture value. The data marketplace owners both sell their products using a usage based model. However, the amount of revenue that they generate from their automotive customers is different. TomTom earned €266 million from their automotive customers with a total revenue of €426 million in 2019 (TT-3b, p. 30). INRIX earned $73 million in 2014 from this customer segment with a total revenue of $150 million (IN-5, p. 34).

The dissimilarities we found within the similar taxonomic characteristics of TomTom and INRIX are summarized in Table 20. The differences concern specific details in their business models. Although there is a difference in data that TomTom and INRIX collect from data sellers, this data remains of similar origin and contributes to their creation of dynamic maps. Further, their usage of varying cloud operators does not change the fact that TomTom and INRIX store data centrally and that their different algorithms are part of the same data processing activity, data analysis. The difference in revenue streams from automotive participants seems to be significant. However, the revenue streams from the automotive customer segment makes up for half of the total revenue of both data marketplace owners. Whether the higher revenue streams of TomTom result in a higher profit than INRIX requires further research into the cost model of both companies, which we did not include in our taxonomy. In short, the dissimilarities in these characteristics are negligible compared to the overall similarities as shown in our taxonomy. Therefore, we group TomTom and INRIX in one business model archetype. The archetype is further explained in section 6.2.

Table 20: Pairwise comparison TomTom and INRIX

Dimension	TomTom	INRIX
Participants	Emphasis on data generated by own hardware users	Emphasis on data from corporate participants
Platform infrastructure	Microsoft Azure Cloud	Microsoft Azure Cloud integrated with Amazon AWS Cloud
Data processing activities	Apply proprietary algorithms developed by TomTom	Apply proprietary algorithms developed by INRIX
Revenue streams	€266 million revenue from automotive participants	$73 million revenue from automotive participants

6.1.2 Dissimilarities between IOTA and Ocean Protocol

IOTA and Ocean Protocol both focus on data trade across all industries. They enable direct data trade between their participants with a decentralized platform infrastructure that runs on DLT. Our taxonomy shows a difference in their revenue streams as IOTA transparently communicates to be funded by donations and Ocean Protocol does not provide information about their revenue streams. In addition, there are subtle differences in characteristics that seem similar in our taxonomy.

In the value creation of their business models, IOTA and Ocean Protocol specify their *domain* in varying terms. IOTA envisions a machine-to-machine economy specified towards the IoT industry (IOTA-4a). Ocean Protocol emphasizes the use of data marketplaces for AI development. Their goal is to improve access to large data volumes in order to improve the accuracy of AI models (OP-3b). Thus, IOTA trades data across industries for the benefit of IoT development and Ocean Protocol aims to further advance AI.

IOTA and Ocean Protocol deploy varying decentralized *platform infrastructures* to deliver value. Ocean Protocol uses blockchain technology to provide secure and immutable transactions. IOTA argues against blockchain technology due to the transaction costs of blockchain. IOTA built their own DLT, called the Tangle. They attribute the main benefit of the Tangle to the absence of fees to carry out transactions: "there are no fees, there is no incentive to centralize resources. It also becomes a lot more decentralized just as a consequence of this architecture" (IOTA-4c). The variance in the application of DLTs results in differing decentralized platform infrastructures.

Variations in the implementation of the *data pricing mechanism* and *payment currency* lead to differences in the value capture between IOTA and Ocean Protocol. Ocean Protocol gives the data seller more freedom than IOTA to choose the pricing mechanism for their data. By listing a range of possible fixed and flexible pricing mechanisms (OP-3a), Ocean Protocol advises their participants in choosing a pricing model for their data, but the final decision is made by the data sellers. At IOTA, data sellers have to set usage-based prices for the streamed data. In addition, the data marketplaces vary in their cryptocurrencies. Ocean Protocol trades in Ocean tokens and IOTA uses the IOTA token. Thus, the distinctions in data pricing mechanisms and payment currencies lead to variation in value capture.

The dissimilarities we found when comparing IOTA and Ocean Protocol are summarized in Table 21. Despite the dissimilarities, the data marketplace owners largely match in one another's business models. Although Ocean Protocol allows their participants to choose any data pricing mechanism, the data sellers choose fixed pricing mechanisms in practice. This results in a fixed price that is set by the data seller at both data marketplaces. Although the tokens of the data marketplaces cannot be used interchangeably, the Ocean and IOTA token are both cryptocurrencies. We can also rationalize the difference in domain, because the development of IoT and AI are supported by one another. This is described in an article from Forbes: “AI needs data, IoT needs intelligence and insights, and both need security and transparent marketplaces” (OP-4b). The implementation of different decentralized platform infrastructures by the data marketplace owners is the biggest dissimilarity. Both data marketplace owners deploy a decentralized platform infrastructure, but the implementation of blockchain technology and the Tangle results in different infrastructures. Because this is the only significant difference, we decide to group IOTA and Ocean Protocol in one business model archetype. The archetype is further explained in section 6.2.

Table 21: Pairwise comparison IOTA and Ocean Protocol

Dimension	**IOTA**	**Ocean Protocol**
Domain	Internet of Things	Artificial Intelligence
Platform infrastructure	Blockchain	Tangle
Data pricing mechanism	The data seller is bound to setting a usage based pricing mechanism	The data seller has the freedom to set any fixed or dynamic pricing mechanism
Revenue stream	Donations	No info
Payment currency	IOTA token	Ocean token

6.1.3 Similarities between HERE and Caruso

HERE and Caruso do not deliver the exact same data service and target participants in different domains. Nevertheless, there are similarities in the different characteristics of the dimensions *participants*, *data service* and *data pricing mechanism*. A comparison of these similarities creates insights about the grouping or separation of these data marketplaces in a business model archetype.

HERE and Caruso have some resemblance within the meta-characteristic value creation. We recognize similarities in their type of *participants* and *data service*. Both data marketplaces are consortium owned and list their founding partners as data sellers. Caruso is owned by multinationals such as TecAlliance, Bosch and Continental (CR-3). Audi, BMW and Daimler are the owners of HERE (HE-4b). The members of the consortia of these data marketplaces are well-known OEMs and tier 1 suppliers who function as data seller participants at the data marketplace. Next, we find similarities in their data service. In addition to the sale of customized maps, HERE trades standardized data as a data brokering service. Caruso’s core business is to trade standardized data and provide a data brokering service. Both data marketplaces publish metadata about the datasets on their platform, which can be browsed by participants to find datasets and request access via APIs. Thus, the data marketplaces are similar in their ownership and their value proposition of standardized data output and improved data access through the data brokering service.

The *data pricing mechanisms* of the data marketplaces also show commonalities. HERE provides pricing schemes which the data seller uses to sell their data and set a data price. Caruso advises their participants in pricing mechanisms, but the data seller decides what price to set for their data. Thus, in both cases the pricing models are set by the data sellers.

The similarities in the dissimilar taxonomic characteristics of HERE and Caruso are summarized in Table 22. These similarities are present because HERE provides two different value propositions. HERE offers a brokering service in addition to their customized map service. If we group HERE and Caruso in one business model archetype, the customized value proposition of HERE could not be represented. Therefore, HERE and Caruso should not be grouped into one business model archetype. HERE also cannot be grouped with TomTom and INRIX, because this would hide their application of the brokering service. Thus, we split HERE and Caruso into two individual business model archetypes. The archetypes are further explained in section 6.2.

Table 22: Pairwise comparison HERE and Caruso

Dimension	Similarities
Participants	Data marketplace shareholders are data sellers at the platform
Data service	Data brokering service improves access to datasets via APIs
Data quality	The data marketplaces do not provide information about data quality reviews
Platform infrastructure	Centralized
Data pricing mechanism	Prices of traded datasets via the brokering service are determined by the data seller

6.2 Business Model Archetypes

Similar business model patterns compose business model archetypes for data marketplaces. In the pairwise comparison of TomTom and INRIX we argued that the dissimilarities found in their business model characteristics do not overrule their similarities. Therefore, TomTom and INRIX are clustered in one business model archetype. Similarly, we cluster IOTA and Ocean Protocol in a business model archetype. HERE and Caruso are too dissimilar in their business model characteristics. Their business models are represented by two separate business model archetypes.

We exclude four dimensions from our taxonomy in the business model archetypes. IOTA and Ocean Protocol have different revenue stream characteristics. Thus, the dimension revenue streams is not included our archetypes. In section 5.1.2 we established that all data marketplaces have similar participants. Therefore, this dimension does not contribute meaningful distinctive characteristics in our archetypes. The characteristics of the data processing activities correlate to the data service that data marketplace owners perform. Data marketplace owners who deliver a customized map service perform all data processing activities and data marketplace owners that perform the data brokering service are limited in their data processing activities. The dimension data processing activities does not add meaningful information in comparison to the data service and is excluded from our archetypes. Last, we exclude the dimension payment currency, because this dimension directly relates to the platform infrastructure. Data marketplace owners with decentralized platform infrastructures who use DLT pay in cryptocurrency. Whether data marketplaces use fiat currency or cryptocurrency is deemed less meaningful for our business model archetypes.

We found patterns in the dimensions domain, data service and output, data quality, privacy, contract, platform access, platform infrastructure and data pricing mechanism. The characteristics of these dimensions show how data marketplace owners capture, deliver and create value in distinctive manners. The *aggregating data marketplace* owner performs data analyses as part of their customized map service to provide aggregated data. This is similar to the data marketplace owner of *the aggregating data marketplace with an additional brokering service* who also offers a customized map service. The customized map service is their core business and in addition they perform a data brokering service to enable standardized data trade between their participants. *The consulting data marketplace* owner performs a data brokering service which distincts itself from the other archetypes, because the service is paired with personal assistance of the data marketplace owner through bilaterally negotiated contracts. *The facilitating data marketplace* owner focuses on their data brokering service that runs on a decentralized platform infrastructure. This results in four business model archetypes that are presented in Table 23. In line with our assumption that each data marketplace has one business model and the rule of mutual exclusivity, each data marketplace matches one of the business model archetypes. We further discuss the archetypes in the subsequent sections.

Table 23: Business model archetypes

Archetype	**Aggregating data marketplace**	**Aggregating data marketplace with additional brokering service**	**Consulting data marketplace**	**Facilitating data marketplace**
Case	*TomTom and INRIX*	*HERE*	*Caruso*	*IOTA and Ocean Protocol*
Orientation	*Hierarchical*	*Mixed hierarchical/market*	*Mixed hierarchical/market*	*Market*
Ownership	*Private*	*Consortium*	*Consortium*	*Independent*
Domain	Location	Location	Automotive	Cross-industry
Data service and data output	Customized map service Aggregated data	Both customized map service and data brokering service Both aggregated data and standardized data	Data brokering service Standardized data	Data brokering service Standardized data
Data quality	Reviews by data marketplace owner	Reviews by data marketplace owner	No info	Reviews by users
Privacy	Anonymized	Encrypted	Anonymized	Encrypted
Contract	Negotiated contract	Both negotiated and standardized contract	Negotiated contract	Standardized contract
Platform access	Closed	Open	Closed	Open
Platform infrastructure	Centralized	Centralized	Centralized	Decentralized
Data pricing mechanism	Set by data marketplace owner	Both set by data marketplace owner or data seller	Set by data seller	Set by data seller

6.2.1 Aggregating Data Marketplace

TomTom and INRIX apply the *aggregating data marketplace* archetype. They create value for their customers by aggregating the data from their data sellers to provide tailored maps for their customers. Through *bilaterally negotiated contracts* the data marketplace owners establish *personal customer relationships* with the data marketplace participants. The data marketplace owners have close contact with their participants during bilateral negotiations to understand and define data trading conditions. They personally assist their customers during data collection. Osterwalder & Pigneur (2010) introduce personal assistance as a manner for business owners to build customer relationships through human interaction. Although the personal interaction demands investment from the data marketplace owner, the creation of personal customer relationships should increase the commitment of customers to their data marketplace. Moreover, the data marketplaces have *well-understood customer segments* in the *location* domain. Within the location domain, the data marketplace owner knows who their participants are, where the data comes from, what information it contains and what purpose the data is used for. Their customer groups are segmented. For example, the automotive segment comprises OEMs and Tier 1 suppliers and the enterprise segment comprehends mobile application developers, cloud providers and fleet managers. Segmented customers have slightly different needs and problems and receive differing value propositions (Osterwalder & Pigneur, 2010). This leads to the *customized value proposition* that the data marketplace owner creates by offering a *customized map service*. The data marketplace owner combines a real-time traffic service, EV service, parking service and speed camera service to create customized maps. As such, an OEM receives aggregated real-time traffic information (RTTI) and parking information for their navigation system while an external developer receives EV data to develop their own charging application. The data quality is assured by the data marketplace owner who reviews and cleans data. The data marketplace owner handles the payments, contracts and provides the infrastructure for all participants to satisfy their needs.

The aggregating data marketplace has *closed platform access*. The data marketplace owner needs to approve data seller or buyer registration before data can be sold or bought from and to the data marketplace. This contributes to a controlled environment to which participants can be denied. Furthermore, the aggregating data marketplaces need a *centralized platform infrastructure*. The centralized platform infrastructure is connected to the customer IT systems and realizes a central access point for the data marketplace owner to modify the data and perform their service.

At data marketplaces of the aggregating data marketplace archetype, *the data marketplace owner sets the price of the traded data*. The aggregated data output is owned and sold by the data marketplace owner. The usage-based data that is sold leads to direct revenue streams for the data marketplace owner.

6.2.2 Aggregating Data Marketplace with Additional Brokering Service

HERE applies the aggregating *data marketplace with an additional brokering service* archetype. This archetype includes two distinct value propositions. One value proposition is similar to the value proposition of the aggregating data marketplace presented in section 6.2.1. Data marketplace owners of both archetypes focus on delivering a customized value proposition and aggregated data within the location domain. However, data marketplace owners who apply the *aggregating data marketplace with an additional brokering service* archetype offer a second, standardized value proposition which is the data brokering service. This service enables standardized data trade directly between data sellers and data buyers at the data marketplace. The data marketplace owner uses negotiated contracts for their customized value proposition and standardized contracts for their standardized value proposition. The standardized contract enables automated assistance. Automated assistance has lower costs than personal assistance and can handle a large number of users (Osterwalder & Pigneur, 2010). The application of both negotiated and standardized contracts enables the data marketplace owner to offer personal assistance to some customers while simultaneously serving many other participants through automated assistance.

The aggregating data marketplace with an additional brokering service has *open platform access*. Anyone who creates a user account can enter the platform. The data marketplace owner deploys a *centralized platform infrastructure*. Similar to the explanation in section 6.2.1, the central storage of data is required for the data marketplace owner to perform data collection, standardization, cleansing, storage, analysis and distribution and deliver the customized value proposition.

To capture value, the data marketplace owner maintains two data pricing mechanisms. The *data marketplace owner sets the price* for the aggregated data that is produced with the customized map service and the *data sellers set the price* for the standardized data that they sell via the brokering service.

6.2.3 Consulting Data Marketplace

Caruso applies the *consulting data marketplace* archetype. They offer a *standardized value proposition*, as does the data marketplace owners who apply the *aggregating data marketplace with additional brokering service* archetype. Significant for the brokering service of the data marketplace with *consulting data marketplace* archetype is that the data marketplace owner pairs the service with negotiated contracts. The data marketplace owner negotiates the contract conditions with their participants bilaterally. The data marketplace owner gains knowledge about the data needs and price preference of their participants and aligns the needs of their data sellers and data buyers. If a data seller wants to sell specific data assets at the marketplace, there needs to be a data buyer interested in buying those segments and vice versa. The participants are personally assisted on bilateral basis by the data marketplace owner through negotiated contracts. Similar to the contracts of the aggregating data marketplace, these contracts lead to *strong customer relationships*. The customers are identified in the automotive domain. The data marketplace owner aims to serve all participants with an interest in automotive data. Potential participants are OEMs, any supplier of the OEM, insurance companies, infotainment services and external developers. The consulting data marketplace is the intermediary who connects these interdependent groups.

The consulting data marketplace has *closed platform access*. Participants may enter the platform after they are provided login credentials by the data marketplace owner. This provides controlled provision and purchase of data at the marketplace. Furthermore, consulting data marketplaces have a *centralized platform infrastructure*. The data marketplace owner stores and publishes metadata about the datasets in the centralized platform infrastructure. The metadata is analyzed to create insights about the platform usage patterns. Significant for the data marketplace with the consulting data marketplace archetype is that the exchanged data sets are not stored in their cloud. Only metadata about the datasets is stored.

The consulting data marketplace allows *the data seller to determine the price* of the sold data. The data marketplace owner consults their participants about possible data pricing mechanisms. The revenue streams for the exchanged data are transferred between the data seller and buyer. The data marketplace owner receives a commission of the sold data from the data seller and is paid for their provided service.

6.2.4 Facilitating Data Marketplace

IOTA and Ocean Protocol apply the *facilitating data marketplace* archetype. They coordinate transactions between data sellers and buyers through the data brokering service without interference of the data marketplace owner. The facilitating data marketplace contains a *standardized value proposition* that comprises a data brokering service. The data marketplace owner aims to provide access to data that participants did not have access to before to further develop IoT and AI technologies. Developers lack data to improve their algorithms and larger companies lack advanced algorithms to analyze their data. At the data marketplace, these participants can trade data across *all industries*. The participants process the standardized data and review the data quality themselves, with minimal interference of the data marketplace owner. The data marketplace owner does not offer personal assistance like data marketplace owners who apply the *consulting data marketplace* archetype, but uses standardized, smart contracts. This foresees a high number of transactions between participants and *automizes the trading process*.

The facilitating data marketplace has *open platform access*. Anyone who knows how to use the infrastructure and has a need to trade data can join the ecosystem. In addition, the facilitating data marketplace is the only business model archetype that includes a *decentralized platform infrastructure*. The DLT is the building block that facilitates the value proposition of the data marketplace owner. The decentralized platform infrastructure allows for minimal intervention of the data marketplace owner and direct transactions between the data seller and buyer. Transactions in DLTs are immutable and transparent, to ensure safe data delivery. The main task of the data marketplace owner is to define transaction rules and link transactions to be executed and verified by the participants.

The marketplace owners who apply the facilitating data marketplace archetype enable *the data sellers to set the price* for the traded datasets. The revenue streams are directly transferred between the data seller and data buyer. These data marketplaces are owned by non-profit organizations. They do not intend to make profit from the data that is traded at their platform.

6.3 Conclusion of Chapter 6

In this chapter a pairwise comparison is performed to support the creation of distinctive business model archetypes. This enables us to answer the third sub-question: *What business model archetypes can be identified for data marketplaces?* Four business model archetypes emerge. These are (i) the *aggregating data marketplace* that includes a customized map service in which the data marketplace owner analyzes data internally to aggregate data into a customized value proposition, (ii) the *aggregating data marketplace with an additional brokering service*, applied by data marketplace owners who perform data aggregation to create a customized value proposition as core business and provide a data brokering service as additional standardized value proposition, (iii) the *consulting data marketplace* that data marketplace owners apply to provide a data brokering service and advise participants in their data usage and exchange and (iv) the *facilitating data marketplace* with a decentralized platform infrastructure deployed by the data marketplace owner to coordinate transactions between data sellers and buyers in their data brokering service.

7. Evaluation

In this chapter, our taxonomy is evaluated on completeness and usefulness. As explained in section 3.5, we interviewed Spiekermann (2019) and Fruhwirth et al. (2020) to evaluate our taxonomy. We consider them experts in the development of business model taxonomies for data marketplaces and incorporate their feedback in this chapter. In section 7.1, we evaluate whether our sample is representative for data marketplaces. Next, we evaluate our taxonomy based on the objective and subjective ending conditions in section 7.2. If all conditions are satisfied, our taxonomy development process is finished. Consequently, we evaluate the usefulness of our refined taxonomy and archetypes in section 7.3. We aggregate the information from these sections in section 7.4 to improve our taxonomy or make recommendations for future research. Finally, the fourth sub-question is answered in section 7.5.

7.1 Evaluation of the Data Marketplace Sample

Within the limited timeframe of this thesis we could not examine all existing data marketplaces. The B2B automotive industry is chosen as industry of focus as explained in section 1.3. From this industry, we classified TomTom, INRIX, HERE, Caruso, IOTA and Ocean Protocol in our taxonomy. These data marketplaces are selected to represent three different data marketplace types: (i) data marketplaces with a hierarchical orientation and private ownership, (ii) data marketplaces with a mixed hierarchical and market orientation and consortium ownership and (iii) data marketplaces with a market orientation and independent ownership. Compared to the 16-20 data marketplaces classified in the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020), our sample of six data marketplaces is small. However, we opted for a small sample to be specific in our interpretation of business model characteristics.

We discussed our sample with Spiekermann and Fruhwirth who understood our decision to include various types of data marketplaces. Both experts named Otonomo to expand our sample. Otonomo is a data marketplace with a mixed hierarchical and market orientation and private ownership. Furthermore, the data marketplaces DAWEX and oneTRANSPORT are suggested as data marketplaces with a market orientation and private ownership. One interviewee explained that DAWEX is not industry specific, but the data marketplace is past the conceptual stage and could be fitting for our sample. The inclusion of Otonomo, DAWEX and oneTRANSPORT would lead a more representative sample of data marketplaces that covers five different data marketplace types. Nonetheless, the experts agreed that our sample is satisfactory to represent three data marketplace types in the B2B automotive industry.

7.2 Assessment of the Objective and Subjective Ending Conditions

According to Nickerson et al. (2013), a taxonomy is finished when it satisfies all objective and subjective ending conditions. Whether our taxonomy meets the objective and subjective ending conditions has been evaluated in two manners. First, we assessed whether our taxonomy met the objective ending conditions in each iteration of our development process. The evaluation of a taxonomy during the building process is known as ex ante evaluation (Oberländer et al., 2019; Szopinski et al., 2019). Similar to Fruhwirth et al. (2020) and Oberländer et al. (2018), we base the assessment of the objective ending conditions on our own judgement. Second, we evaluate whether our taxonomy meets the subjective ending conditions. The subjective ending conditions leave room for interpretation. Therefore, they are assessed based on the expert feedback.

Seven iterations were required to meet all objective ending conditions (see Table 24). Throughout these iterations, we consulted various data sources to build our taxonomy. During the first iteration, we based the outline of our taxonomy on the Business Model Canvas of Osterwalder & Pigneur (2010) and STOF model of Bouwman et al. (2008). In the second and third iteration, we derived dimensions and characteristics from interviews with data marketplace owners and from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). In the fourth, fifth and sixth iteration, we derived dimensions and characteristics from case documents of TomTom, INRIX, HERE, Caruso, IOTA and Ocean Protocol. During the seventh iteration, we revised the case documents and read news releases without identifying new dimensions or characteristics. Thus we reached saturation in the seventh iteration.

Table 24: Assessment objective ending conditions

Objective ending Conditions		Iterations 1	2	3	4	5	6	7	Rationale
OE1	All objects or a representative sample of objects have been examined						x	x	We satisfied OE1 after all cases from our sample were classified in iteration 6.
OE2	No object was merged with a similar object or split into multiple objects in the last iteration	x	x	x	x	x	x	x	As we classified data marketplaces with the assumption that each data marketplace has one business model, no objects were merged or split.
OE3	At least one object is classified under every characteristic of every dimension							x	Characteristics were only removed after we classified all objects. Up to iteration 7, conceptual characteristics such as real-time pricing mechanisms were part of our taxonomy
OE4	No new dimensions or characteristics were added in the last iteration							x	New characteristics of IOTA and Ocean Protocol were still added in iteration 6. For example, IOTA receives donations as revenue stream which other data marketplace do not.
OE5	No dimensions or characteristics were merged or split in the last iteration	x	x	x		x	x	x	A dimensions was split once In iteration 4 we split the dimension *data product* into *data output* and *data service* as explained in section 3.3.3.
OE6	Every dimension is unique and not repeated (i.e., there is no dimension duplication)	x	x				x	x	We removed the dimension *data source* from our taxonomy in iteration 6 because it overlapped with the dimension *participants* as explained in section 3.3.3. From iteration 3 to 5, the dimension data source was part of our taxonomy.
OE7	Every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension)	x	x	x	x	x	x	x	The characteristics within the dimensions were unique in each iteration

The subjective ending conditions are evaluated with experts in an eighth iteration. They assessed our taxonomy as extendible and explanatory. The conciseness, robustness and comprehensiveness of our taxonomy could be improved.

- Both experts indicated that the number of dimensions and characteristics provides a clear overview, but the conciseness can be further improved. One expert suggested to remove the dimension *data processing activities* or *data service* because their characteristics overlap.
- The robustness of the taxonomy is dimension dependent. One expert assessed the dimensions *privacy*, *platform access* and *platform infrastructure* as robust. The characteristics of these dimensions differentiate objects clearly. Other dimensions, such as *domain*, have grey areas that make the distinction of objects difficult. For example, data marketplaces that do not focus solely on location and do not cover the complete automotive domain cannot be categorized easily
- The robustness of the taxonomy changes as data marketplaces further develop. It must be clear that the taxonomy designed in this research provides a current overview of the business models of data marketplaces. One expert elaborated: "*maybe in the current state of this domain the taxonomy is robust, but in the future development of data marketplaces there could be more characteristics*". Because data marketplaces develop quickly our taxonomy needs to be revised in the future to stay robust.
- There are data marketplaces that cannot be classified in the characteristics of the dimension revenue streams. For example, a data marketplace with a revenue stream combination of flat fee tariff and usage based could not be classified in our taxonomy. A new characteristic for each available combination of revenue streams causes the taxonomy to be less comprehensive.

Based on the feedback of the experts, we adapt two elements in our taxonomy to improve its conciseness and comprehensiveness. The dimension *data processing activities* is removed because it overlaps with the dimension *data service*. The dimension *data service* includes services such as purchasing and contract management which are not included in the dimensions *data processing activities*. To retain the dimension with the most concise information in our taxonomy, we remove *data processing activities*. Second, we change the name of the characteristic *no info* to *other* in the dimension *revenue streams*. This enables the classification of data marketplaces with diverse revenue streams without complicating the comprehensiveness of our taxonomy. These adaptions result in a more concise and comprehensive taxonomy. The robustness of our taxonomy is satisfactory for present time, but requires to be revised in the future.

7.3 Usage of Our Taxonomy and Archetypes

Our objective in this thesis is to clarify the business models of data marketplaces in the B2B automotive industry. We primarily developed our taxonomy and archetypes to achieve this objective. Nonetheless, researchers and practitioners could use our taxonomy as well. The experts expressed that they can use our taxonomy to classify and compare data marketplaces. Especially while the definition of a data marketplace is not clear, our taxonomy is useful to determine what characteristics a platform must have to be a data marketplace. Moreover, practitioners may use our taxonomy for a competitor analysis. By mapping competitors in the taxonomy, data marketplace owners could observe what strategic decisions their competitors make. Such observations can motivate them to make adaptions in their own business model to either compete in a similar market as their competitors or diverge to another market. Similar use cases are envisioned for the archetypes. One expert expressed that archetypes may be more suitable to show a complete business model in a comprehensive manner and the taxonomy could be more fitting to compare specific characteristics. Overall, he thought both artefacts would be useful to distinguish business models of data marketplaces and perform a competitor analysis.

7.4 Incorporation of the Evaluation Remarks in this Thesis

Based on the remarks in the previous sections, we improve our taxonomy and make suggestions for future research. The remarks are summarized in Table 25.

Table 25: Evaluation incorporation

Remark	Explanation	Thesis incorporation
Remove data processing activities or data service	The data service and data processing activities overlap. Data marketplace owners who deliver a customized map service perform all data processing activities. Offering the data brokering service entails performance of a limited number of data processing activities by the data marketplace owner. Therefore, these dimensions overlap in their characteristics.	*Adaption in the evaluated taxonomy*. Because the dimension data service includes additional aspects, the dimension data processing activities is removed from our taxonomy (see appendix H)
Define the characteristics of domain more granularly	Location is one domain within the automotive industry. In addition to location, more domains can be specified, such as car development, in-use car data and vehicle maintenance. The inclusion of such domains makes the taxonomy more robust and reliable	*Suggestion for future research*. In the selection of data marketplaces in this thesis, singular domains such as car development and sales do not appear. To identify those domains, more data marketplaces have to be included that are industry specific. If the domains are not identified in practice, data buyers and sellers could be surveyed to gain better insight in the data needs of the participants
Change a characteristic in the dimension revenue streams	The characteristics of revenue streams make the taxonomy less comprehensive. Multiple combinations of revenue streams can be present, that results in a long sum of characteristics. To avoid an uncomprehensive taxonomy, one characteristic is changed to “other” in which data marketplaces with various revenue streams can be classified.	*Adaption in the evaluated taxonomy*. The characteristic that is named “no info” is changed to “other”. Data marketplaces that have a combination of revenue streams other than currently shown in the taxonomy can be classified in the characteristic “other”.

Use the taxonomy and archetypes for research	We use our taxonomy to identify the business models of six data marketplaces. The classification of additional data marketplaces helps researchers to recognize more characteristics and better define what data marketplaces are. For research purpose, it is of interest to classify conceptual data marketplaces as well, because data marketplaces develop fast. For example, dynamic pricing mechanisms are expected to advance the data marketplaces with a market orientation.	*Suggestion for future research*. Add more data marketplaces from various industries. An interest of researchers is to see how data marketplaces can advance in the future. Therefore, conceptual data marketplaces are allowed to be classified to research components that are not applied in practice yet.
Use the taxonomy and archetypes for practice	It is speculated that our taxonomy is useful for practitioners. Practitioners could use the taxonomy to map their competitors and use the taxonomy as a blueprint for their own business model. However, it is unsure whether our taxonomy is truly useful for practitioners to gain new insights about their own business model or the business model of competitors	*Suggestion for future research*. Study how practitioners would use a taxonomy and design a taxonomy for that purpose. We used our taxonomy for research purpose. The usefulness of a business model taxonomy for practitioners requires observation over time

7.5 Conclusion of Chapter 7

In this chapter we assessed the objective and subjective ending conditions and explored the usage of our taxonomy and archetypes to answer the fourth sub-question: *Is the business model taxonomy developed in this research evaluated as complete and useful?* Our taxonomy is based on a sample that is satisfactory to represent three types of data marketplaces in the B2B automotive industry. The taxonomy can still be extended to cover more types of data marketplaces. For example, Otonomo, DAWEX and oneTransport are identified as data marketplaces with new combinations of orientation and ownership. The classification of these data marketplaces is recommended for future research to improve the reliability of our taxonomy. After eight iterations of inducing and deducing dimensions and characteristics from various sources, we satisfied all objective and subjective ending conditions. Experts could not identify characteristics or dimensions that are missing from our taxonomy. We incorporated their suggestions to improve the conciseness and comprehensiveness of our taxonomy, after which we evaluated our taxonomy as complete.

Furthermore, our taxonomy can be useful for researchers and practitioners. Researchers may use our taxonomy to further define what a data marketplace is. The experts understood that the inclusion of data marketplaces with a hierarchical orientation and private ownership enables research of data marketplaces in practice. In defining what a data marketplace is, researchers can use our taxonomy to decide about the characteristics that should be part of a data marketplace. Our taxonomy is useful to identify the business model characteristics that data marketplace owners currently apply. It cannot be used to identify theoretical concepts for future development. Practitioners may use our taxonomy to make design choices for their own business model and do a competitor analysis. We did not evaluate whether practitioners would actually use our taxonomy for such purposes and recommend this for future research.

8 Discussion

In this thesis, we found that owners of differing data marketplace types apply distinctive business model archetypes (see Figure 19). TomTom and INRIX, the data marketplace types with private ownership and a hierarchical orientation apply the *aggregating data marketplace* archetype. HERE and Caruso, data marketplace types with consortium ownership and characteristics from both a hierarchical and market orientation apply the archetypes *aggregating data marketplace with additional brokering service* and *consulting data marketplace*, respectively. IOTA and Ocean Protocol, data marketplace types with independent ownership and a market orientation apply the *facilitating data marketplace* archetype.

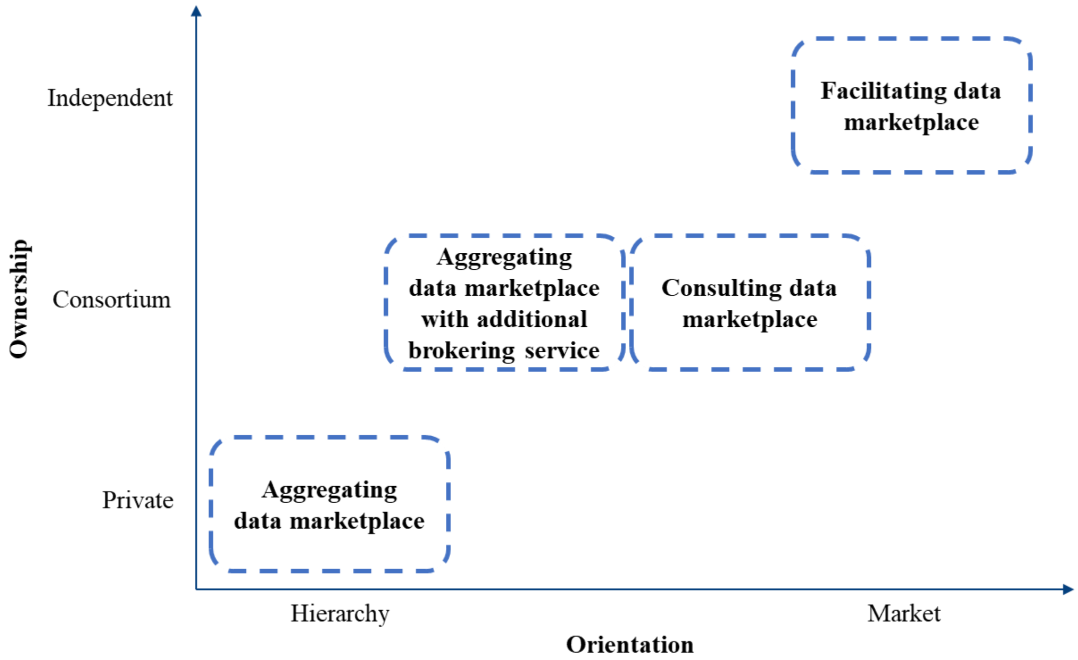

Figure 19: Business model archetypes corresponding to data marketplace orientation and ownership

In the following sections these findings are placed in current scientific literature. In section 8.1 we discuss the value proposition of the data marketplace owners and find that the value proposition of the facilitating data marketplace archetype may not solve their customer problem. In section 8.2 we discuss the establishment of customer relationships by data marketplace owners and find that the data marketplace owners who apply the facilitating data marketplace archetype could fail to attract customers. Next, in section 8.3 we discuss whether we achieved a more specific interpretation of business model components in our taxonomy of business models for data marketplaces in the B2B automotive industry compared to the more general taxonomies by Spiekermann (2019) and Fruhwirth et al. (2020). In section 8.4, contradicting business model characteristics of our taxonomy are compared to the taxonomies developed thus far. Subsequently, we challenge our definition of data marketplace types in section 8.5. Finally, we address the research limitations in section 8.6.

8.1 A Value Proposition that Offers a Solution Instead of an 'Item'

In its essence, all data marketplaces trade data. By performing additional services in their value proposition, a data marketplace owner distinguishes their marketplace from other data marketplaces. Our business model archetypes show that data marketplace owners create additional value for their customers by performing a customized map service, reviewing the data quality or offering personal assistance through negotiated contracts. The value proposition of data marketplaces with the *facilitating data marketplace* archetype is the only value proposition that focuses solely on a data brokering service.

The value proposition of the *facilitating data marketplace* represents the problem that Teece (2010) describes as the sale of 'items' instead of the sale of a solution. Data assets, or 'items', could be described as ‘intangibles’, ‘know-how’ and ‘technological components’. These goods are difficult to price and are rarely traded in market structures (Koutroumpis et al., 2017; Powell, 1990; Teece, 2010). According to Teece (2010), it is a common problem that the sale of assets that do not have perfect property rights, leads to market failure. Business owners who apply business models that are based on selling intangibles may not capture significant value with their value proposition. Therefore, companies who trade intangible assets need to bundle them into a solution.

The *aggregating data marketplace, aggregating data marketplace with additional brokering service* and *consulting data marketplace* archetypes comprise value propositions in which data is bundled into a solution. The data marketplace owners of these archetypes trade data and provide complementary services such as a customized map service, data quality reviews or personal consultation about data sale and purchase. Spiekermann (2019) argues that the performance of such services is a key success factor in the business model of data marketplaces because it increases value for the customer. He finds that data marketplace owners who aggregate data or assure data quality, create value as they go beyond data forwarding (Spiekermann, 2019). The performance of such services does require higher investment in time and money from the data marketplace owner. Osterwalder & Pigneur (2010) explain that companies that focus on value creation deliver a premium value proposition and have a value-driven business model. Their customers do not only pay for the product that they get, but also for the service that the company performs. Data marketplace owners who do not only trade in data assets, but also deliver a value-adding service can sell their solution against a higher price, which their customers are willing to pay for.

Data marketplace owners who apply the *facilitating data marketplace* archetype focus on data forwarding with their brokering service. These data marketplace owners have a lean cost structure and automize most of their processes. This is what Osterwalder & Pigneur (2010) describe as a cost-driven business model. Data marketplace owners who apply this business model promise an increase in data accessibility for their participants against a low price. Their value proposition entails trade in data ‘items’. However, this does not appear to be the solution for their customers. Data sellers and buyers remain absent, which diminishes their ability to increase data access. There is a need for data marketplace owners who apply the *facilitating data marketplace* archetype to bundle their data brokering service with complementary services. In this way they can attract data sellers or buyers by offering a solution instead of trading data ‘items’.

8.2 The Establishment of Personal Customer Relationships or Competitive Pricing

The relationships that are created with participants differ between data marketplaces with a hierarchical orientation compared to data marketplaces with a market orientation. Owners of data marketplaces with a hierarchical orientation form personal customer relationships which leads to returning customers. Data marketplaces with a market orientation do not form personal customer relationships and need competitive pricing to attract customers.

We observed that owners of data marketplaces with a hierarchical orientation or mixed hierarchical and market orientation apply the *aggregating data marketplace*, *aggregating data marketplace with additional brokering service* and *consulting data marketplace* archetypes. These archetypes have negotiated contracts. Data marketplace owners who apply these contracts trade data bilaterally and form personal customer relationships. They offer personal assistance to their customers regarding data sale and purchase. Koutroumpis et al. (2020) call negotiated contracts in hierarchical structures “relational contracts” that are long term and enable repeated interaction between the data marketplace owner and their participants (Koutroumpis et al., 2020). We expected negotiated contracts and the formation of personal customer relationships in organizations with a hierarchical trading structure. As Powell (1990) explains, the personal identification between the trading parties in a hierarchy causes them to trade repeatedly with each other. Actors who trade in these organizations are driven by routines and have less room to display opportunistic behavior (Powell, 1990; Williamson, 1973). This causes participants to return to the data marketplace.

Owners of data marketplaces with a market orientation do not form personal relationships with their customers. These data marketplace owners apply the *facilitating data marketplace* archetype. They have standardized contracts to offer automated assistance and lower transaction costs. From chapter 4.1 we recall that automated assistance enables low operational costs and high transaction speed. These are desired conditions for organizations with a market structure where trading parties seek quick and efficient interactions. In market structures, buyers minimize their personal costs and easily switch between sellers when they are not satisfied by certain pricing conditions (Williamson, 1973). Powell (1990) confirms that price competition highly influences the behavior of actors in markets. To satisfy the needs of their participants, owners of data marketplaces with a market orientation need to set a competitive environment and keep product prices low.

This requires dynamic pricing mechanisms and high numbers in demand and supply. However, the high number of data sellers and data buyers has not yet been reached at data marketplaces with a market orientation. These data marketplace types fail or remain in the conceptual phase (Koutroumpis et al., 2017; Spiekermann, 2019). As shown in our business model archetypes, dynamic pricing mechanisms do not occur either. Instead, fixed data pricing mechanisms, set by the data marketplace owner or data seller, are applied in practice. Fruhwirth et al. (2020), who researched 20 data marketplaces, found that 2 data marketplace owners establish prices based on auction or negotiation. The other data marketplaces they researched have fixed pricing mechanisms. Out of the 16 data marketplaces that Spiekermann (2019) researched, only 4 data marketplace owners price data based on market supply and demand. One of those data marketplaces withdrew from the market and the others are still in the conceptual stage. The expected functioning of the invisible hand of the market remains obsolete. Because a competitive environment is not established, data marketplaces with a market orientation fail to attract participants who trade on competitive basis and aim for the maximum individual gain.

8.3 Specificity in Business Model Components

One of our identified research gaps in section 1.4 was current taxonomies show a general interpretation of business model dimensions and characteristics. We attributed this to the focus of Spiekermann (2019) and Fruhwirth et al. (2020) on cross-industry data marketplaces. We expected that we would generate more specific business model characteristics in a taxonomy based on data marketplaces in the automotive industry. However, the discovery of industry specific business model characteristics in our taxonomy leaves room for improvement.

Despite our intention to only include data marketplaces that specifically focus on the automotive industry, we include two data marketplaces that focus on multiple industries; IOTA and Ocean Protocol. The inclusion of these data marketplaces interferes with our goal of to find specific business model characteristics in one industry. We selected these cross-industry data marketplaces, because we could not identify data marketplaces with a market orientation and independent ownership that focus on the automotive industry only. Additional exploration is required to find these data marketplaces. The analysis of additional data marketplaces could extend our taxonomy with more specific characteristics. However, data marketplaces with a market orientation and independent ownership that focus only on the automotive industry may not exist. According to Koutroumpis et al. (2020), multilateral data marketplaces target participants across the data ecosystem. Data marketplace owners from these types of data marketplaces may have no interest in focusing on one industry only and rather target multiple industries to increase the number of potential participants at their platform.

We expected to find industry specific characteristics for the B2B automotive industry in the dimensions *domain* and *participants.* In the dimension *domain*, only the characteristic *location* specifies the automotive industry on a more granular level. This domain is targeted by TomTom, INRIX and HERE. We expected to find additional domain characteristics such as car development, in-use car data and vehicle maintenance. We identified such characteristics for Caruso who defines domains such as vehicle health, vehicle non-powertrain hardware and vehicle powertrain resources. However, a sequence of many characteristics can be overwhelming, which is undesirable for the conciseness of a taxonomy (Nickerson et al., 2013). Therefore, we aggregated these characteristics into one which is called the *automotive* domain in our taxonomy. Similar reasoning is applicable to the dimension *participants*. In section 5.1.2 we explained that all data marketplace owners list participants such as OEMs, Tier 1 suppliers, insurance companies, fleet managers, broadband providers, software developers, etc. To distinguish the characteristics of the data marketplace participants, we would have to present company brands in our taxonomy. This would undermine the conciseness of our taxonomy. We made a trade-off between the conciseness of our taxonomy and the granularity in business model characteristics. This contributed to a less specific business model taxonomy for the B2B automotive industry.

8.4 Contradicting Business Model Characteristics

The data marketplaces HERE, Caruso and IOTA are classified in our taxonomy as well as in the taxonomy developed by Spiekermann (2019). Overall, the identified characteristics of these cases are similar in our taxonomies. However, we specified three characteristics in the business models of IOTA and Caruso differently in our taxonomy compared to the taxonomy of Spiekermann (2019). Spiekermann (2019) defines the pricing mechanism of IOTA as dynamic, whereas we identified the pricing mechanism of IOTA as fixed because data prices are set by the data seller. This difference can be explained by the fact that Spiekermann (2019) includes conceptual characteristics in his taxonomy. IOTA envisions pricing mechanisms that are dynamically set in the real-time market. However, in practice the data prices at IOTA are fixed and set by the data seller. Furthermore, in the taxonomy of Spiekermann (2019) compared to our taxonomy, different revenue models were identified for both IOTA and Caruso. Spiekermann (2019) states that IOTA generates revenue from transaction fees, while we note that IOTA receives revenue from donations. Spiekermann (2019) also identifies a different revenue model for Caruso. According to Spiekermann (2019), Caruso charges a membership fee, whereas we found that Caruso takes a margin of the data that that their data sellers sell. These deviations can be explained by the changes in revenue model that occur over time of both data marketplace owners. As recognized by Spiekermann (2019) and Fruhwirth et al. (2020), data marketplaces undergo rapid change and their business model characteristics are likely to change. Therefore, business model taxonomies of data marketplaces need to be revised and adapted over time.

8.5 Challenging Data Marketplace Types

There is no unanimous definition of a data marketplace in literature yet. Researchers have different perspectives on what a data marketplace entails and how it should function. This causes our definition of data marketplace types to be vulnerable to critique.

On the one hand, researchers can criticize our definition of data marketplace types to be too broad, as in this research we define both hierarchical and market oriented structures of data marketplaces. According to Schomm et al. (2013) and Spiekermann (2019), data marketplaces enable sellers and buyers to exchange data among themselves at a multi-sided platform. They exclude organizations with a hierarchical orientation and private ownership, that buy data from their data sellers, aggregate data and sell data to data buyers, from their definition of a data marketplace. According to Schomm et al. (2013) and Spiekermann (2019), such organizations are data vendors. They would consider TomTom and INRIX data vendors instead of data marketplaces. Scholars who adopt this view could argue that only multi-sided platforms with a market orientation are considered data marketplaces.

On the other hand, our definition of data marketplace types can be criticized to be too narrow because we did not include the network orientation in our definition of a data marketplace. According to Powell (1990), network organizations are neither markets nor hierarchies. As explained in section 2.2.2, organizations with a network orientation trade assets whose value is difficult to measure on collaborative basis between actors who aim to achieve mutual benefit. Data could be suitable for trading at data marketplaces with a network orientation, as their value is difficult to measure (Koutroumpis et al., 2017; Powell, 1990). Powell (1990) insists network structures are trading structures that need to be included in addition to hierarchies and markets to provide a satisfactory overview of economic exchange.

When taking a closer look at data marketplaces that are owned by consortia (Caruso and HERE), their trading structure may be closer to a network orientation than to a hierarchical or market orientation. We identified Caruso and HERE as marketplaces with a mixed hierarchical and market orientation. According to Stahl et al. (2016), consortia-based marketplaces are marketplaces led by a collaboration of companies in the same industry who aim to improve their processes. Moreover, Koutroumpis et al. (2017) explain that actors in data consortia have shared interest to trade data. Collaboration and shared interest are typical for organizations with a network structure (Powell, 1990). Furthermore, Koutroumpis et al. (2017) describe that consortia often maintain closed platform access to restrict entry for companies who are not in the consortium. We observed closed access in the *consulting data marketplace* archetype that is applied by Caruso. Their participants are either part of the consortium or trusted partners of the consortium. HERE is consortium owned as well and applies the *aggregating data marketplace with additional brokering service* archetype. They have open platform access to foster collaboration and data sharing among participants. These observations show that not all data marketplaces in practice can be classified as data marketplaces with a hierarchical orientation on one end and a market orientation on the other end. Consortium owned data marketplaces might be better described as data marketplaces with a network orientation.

Though hierarchical oriented marketplaces are not widely accepted as data marketplaces, we deliberately included these data marketplace types to allow research of data marketplaces in practice. As explained in section 2.2, data marketplaces in a network context were not researched, because we observed a market-hierarchy continuum in practice and literature. However, as explained above, data marketplaces with consortium ownership may better fit the description of an organization with a network orientation. Therefore, the hierarchy – market continuum is too narrow to research data marketplaces. This is in line with Powell (1990) who calls for the consideration of networks as a third trading structure. Research about data marketplaces in a network orientation would raise new questions. Do data marketplaces with a network orientation put more focus on facilitating collaboration or co-creation among businesses? Is this something that data marketplace participants value? What service should the data marketplace owner provide in a network orientation? Would data marketplaces with a network orientation have open or closed platform access? We leave these questions for future research.

8.6 Research Limitations

In section 3.6, five methodological limitations are mentioned that influence the results. First, we were aware of existing business model frameworks and taxonomies that could have influenced the derivation of new business model dimensions. Out of the five dimensions that we derived based on the Grounded Theory approach, the dimensions *contract* and *data processing activities* are new. The remaining dimensions *platform infrastructure*, *revenue streams* and *data pricing mechanism* were already included in the developed taxonomies thus far (Fruhwirth et al., 2020; Spiekermann, 2019). It is possible that other business model dimensions are not recognized, because we already formed a preconceived opinion about the business model dimensions that should be in our taxonomy. This risk can be decreased by a second researcher who analyzes the interview scripts and compares their codes to our codes.

Next, the selected cases and the researched case sources are limited in numbers. Business model dimensions and characteristics may remain undiscovered in our taxonomy. To improve the robustness of the taxonomy, additional cases and sources could be researched. Nevertheless, we reached theoretical saturation during our last iteration step of our taxonomy development process. This indicates that no new dimensions or characteristics could be derived with the resources that were available. In addition, experts who were interviewed to evaluate our taxonomy agreed that the taxonomy is complete and useful.

Furthermore, the pairwise comparison that was performed to create the business model archetypes focused on similarities and dissimilarities within groups. Based on this comparison, we generated four business model archetypes. A comparison of intergroup differences between the archetypes was not performed. Such an analysis would improve certainty in across-group difference (Eisenhardt, 1989).

Moreover, the number of interviews that were performed to evaluate our taxonomy are limited. We conducted two expert interviews with the researchers who developed business model taxonomies for data marketplaces. This created the advantage that they were familiar with taxonomies and business models of data marketplaces. However, it is possible that the experts were be biased towards their own taxonomies. Although the experts were able to evaluate the taxonomy and suggest points for improvement, they suggested inclusion of dimensions and characteristics that are present in their own taxonomies. This shows that they are biased towards their own taxonomies. Evaluation interviews with experts who did not make a taxonomy before could generate unbiased feedback. Practitioners could for example be asked to classify their own data marketplace in our taxonomy. If they are able to classify their data marketplace in a characteristic of each dimension, our taxonomy could be identified as complete and useful for this user group.

Last, this research is conducted by one person. This presents the risk of misinterpretation of information. Eisenhardt (1989) advises to perform qualitative research in multidisciplinary groups. Additional perspectives on the analyses and results contribute to the generation of new ideas and trustworthiness of the outcome. The development of new taxonomies and the comparison of the results with our taxonomy contribute to the trustworthiness of the outcome and increases understanding of business models of data marketplaces.

9 Conclusion

Companies are increasingly reliant on internal and external data to further advance their business. Currently, businesses mainly produce data for their own usage and store it in data silos afterwards. This hinders the secondary usage of data, when companies reuse external data. Data marketplaces enable secondary data usage. Data marketplaces with a market orientation are envisioned to advance large scale multilateral data trade among companies, but these data marketplace types remain conceptual. In practice, buyers and sellers rather trade data bilaterally at data marketplaces with a hierarchical orientation. Little is known about the business models that data marketplace owners apply to transform conceptual ideas into working value propositions. In literature, researchers focus on one type of data marketplace with a market orientation and independent ownership. This is the data marketplace type that remains conceptual. We analyzed the business models of various types of data marketplaces that range from hierarchical to market orientation and private to independent ownership in the B2B automotive industry. The research question that we address states: *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?*

In our definition, a data marketplace is an organization with a hierarchical or market orientation and private, consortium or independent ownership that matches buyers and sellers, facilitates transactions and provides an institutional infrastructure to trade machine-readable data. We specified the business models of TomTom, INRIX, HERE, Caruso, IOTA and Ocean Protocol in our taxonomy to distinguish one business model from the other. Based on patterns in our taxonomy, we clustered data marketplaces that share similar business model characteristics. This results in four business model archetypes that show distinctive characteristics in the business models of the data marketplaces.

TomTom and INRIX are data marketplaces with private ownership and a hierarchical orientation. They apply the **aggregating data marketplace** archetype. These data marketplace owners process data from their sellers to aggregate data into a customized value proposition. HERE is a data marketplace with consortium ownership and characteristics from both the hierarchical and market orientation. They apply the **aggregating data marketplace with additional brokering service** archetype. HERE aggregates data to create a customized value proposition as core business and provides an additional data brokering service as standardized value proposition. Caruso belongs to the same data marketplace type as HERE and applies the **consulting data marketplace** archetype. They provide a data brokering service and advise their participants about the usage and exchange of their data. IOTA and Ocean Protocol are data marketplaces with independent ownership and a market orientation. They apply the **facilitating data marketplace** archetype. These data marketplace owners deploy a decentralized platform infrastructure to coordinate transactions between data sellers and buyers with their data brokering service.

The owners of data marketplaces with a market orientation and independent ownership, which are conceptual data marketplaces, apply the facilitating data marketplace archetype. This archetype is not proven effective in practice yet. The other business model archetypes are applied by owners of data marketplaces that are past the conceptual stage. The aggregating data marketplace, aggregating data marketplace with additional brokering service and consulting data marketplace archetypes are effective in practice.

9.1 Contribution and Recommendations to Academics

In the emerging research field of data marketplaces, few taxonomies are developed to structure business models of data marketplaces. The taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) cover one type of data marketplace; data marketplaces with a market orientation and independent ownership. These data marketplaces are conceptual. We contribute to academic research by including data marketplace types ranging from hierarchical to market orientation and private to independent ownership. The inclusion of these data marketplace types enables the identification of business models that data marketplace owners actually apply in practice. Our definition of data marketplaces in a market–

hierarchy continuum may be controversial as some researchers would argue that all data marketplaces must have a market orientation. However, we deemed it necessary to include the hierarchical orientation to research business models of data marketplaces in practice. In fact, we recommend future researchers to also include data marketplaces with a network orientation in their research, because data marketplaces with a consortium ownership may be network oriented. Compared to data marketplaces with a hierarchical or market orientation, we expect data marketplaces with a network orientation to have a stronger focus on collaboration among participants. Research about business models of data marketplaces with a network orientation can generate new insights in the role of the data marketplace owner and their value creation, delivery and capture mechanisms.

Our taxonomy offers a starting point for other researchers to further structure business models of data marketplaces and identify new business model dimensions and characteristics. Our taxonomy, based on the B2B automotive industry, can be extended in two ways. On the one hand, additional data marketplaces from the B2B automotive industry can be classified. During the evaluation interviews Otonomo and oneTRANSPORT were suggested as additional data marketplaces in the B2B automotive industry. The classification of these additional data marketplaces may result in a more reliable and exhaustive taxonomy for the automotive industry. On the other hand, more data marketplaces from industries different than the automotive industry can be classified. The insurance industry is for example mentioned by Koutroumpis et al. (2017) as an industry with data marketplaces past the conceptual phase. However, the classification of data marketplaces from other industries may require a more generic interpretation of business model characteristics to enable comparison among cases. The researcher has to make a trade-off between the conciseness of the taxonomy and the granularity in business model characteristics

Furthermore, we developed a research method that can be adopted by future researchers to develop taxonomies. We extended the taxonomy development method by Nickerson et al. (2013) with Grounded Theory, content analysis, cross-case analysis and semi-structured expert interviews. These methods offer guidelines on how to induce and deduce concepts, derive patterns from the taxonomy and evaluate the results. If the number of classified data marketplaces is sufficient, a quantitative clustering method can be applied instead of a cross-case analysis to generate business model archetypes. Quantitative clustering is an objective method to generate archetypes. A comparison of archetypes based on a quantitative method with archetypes based on a qualitative method increases validity of the results.

We did not evaluate our taxonomy with data marketplace practitioners due to time constraints. In cooperation with the experts who designed the existing taxonomies we speculated that our taxonomy could be useful for data marketplace owners. Data marketplace owners can make design choices based on our taxonomy to develop their data marketplaces. For instance, practitioners who are still designing their data marketplace can use our taxonomy to select characteristics for their own business model. Furthermore, practitioners from data marketplaces that are past the conceptual stage can use our taxonomy for a competitor analysis. They can identify whether their competitors are innovating their business models in areas where they should evolve as well. Whether our taxonomy suffices for those purposes is not evaluated and is advised for future research.

9.2 Contribution and Recommendations to Practice

The increase in data generation and the interest of businesses in external data creates opportunities for data marketplace owners. To set up and run a data marketplace, practitioners must understand how to create, deliver and capture value with their business model. Our taxonomy and archetypes offer insights in applicable business model components for data marketplace owners.

Data marketplaces with a market orientation and independent ownership try to enter the market, but have not yet been successful. Owners of these data marketplaces apply the facilitating data marketplace archetype. They focus on a data brokering service to improve data access for their participants. However, they fail to attract participants because data is not offered for competitive prices and the data marketplace owner creates limited value for their customers. The aggregating data marketplace, aggregating data marketplace with additional brokering service and consulting data marketplace archetypes are applied by owners of data marketplaces past the conceptual stage. Data marketplace owners who apply these archetypes offer a solution to their customers by performing a customized map service, reviewing and ensuring the data quality or personally consulting their participants about data sale and purchase. These additional services contribute to their success. Therefore, we advise data marketplace owners who apply the facilitating data marketplace archetype to complement their data brokering service with additional services to create value and attract more customers.

References

Agahari, W., de Reuver, M., & Fiebig, T. (2019). *Understanding how privacy-preserving technologies transform data marketplace platforms and ecosystems: The case of multi-party computation*.

Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. *European Journal of Information Systems*, *19*(3), 359–376. https://doi.org/10.1057/ejis.2010.21

Allee, V. (2008). Value network analysis and value conversion of tangible and intangible assets. *Journal of Intellectual Capital*, *9*(1), 5–24. https://doi.org/10.1108/14691930810845777

Allianz. (2020). *Full Year and Quarterly Earnings Release 2019*. https://www.allianz.com/en/press/news/financials/business_results/200221_Allianz-4Q-FY2019-earnings-release.html

Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic Management Journal*, *22*(6–7), 493–520. https://doi.org/10.1002/smj.187

Bakos, Y. (1998). The Emerging Role of Electronic Marketplaces on the Internet. *Communications of the ACM*, *41*(8), 35–42. https://doi.org/10.1145/280324.280330

Bapna, R., Goes, P., Gupta, A., & Jin, Y. (2004). User heterogeneity and its impact on electronic auction market design: An empirical exploration. *MIS Quarterly*, *28*(1), 21–43.

Bertoncello, M., Camplone, G., Gao, P., Kaas, H.-W., Mohr, D., Möller, T., & Wee, D. (2016). Monetizing Car Data - New Service Business Opportunities to Create New Customer Benefits. In *McKinsey & Company*. https://doi.org/10.1016/j.bushor.2017.07.011

BMW. (2019). *Annual Report 2019*. https://annualreport.bmwgroup.com/2019/

Bosch. (2019). *Annual Report 2019 - Innovation for Times of Transition*.

Bouwman, H., Faber, E., Haaker, T., Kijl, B., & De Reuver, M. (2008). Conceptualizing the STOF Model. In: Bouwman H., De Vos H., Haaker T. (eds) Mobile Service Innovation and Business Models. In *Springer, Berlin, Heidelberg*. https://doi.org/10.1007/978-3-540-79238-3

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. Sage.

Cheng, B., Longo, S., Cirillo, F., Bauer, M., & Kovacs, E. (2015). Building a Big Data Platform for Smart Cities: Experience and Lessons from Santander. *Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress 2015*, 592–599. https://doi.org/10.1109/BigDataCongress.2015.91

Chesbrough, H. (2010). Business model innovation: Opportunities and barriers. *Long Range Planning*, *43*(2–3), 354–363. https://doi.org/10.1016/j.lrp.2009.07.010

Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. *Industrial and Corporate Change*, *11*(3), 529–555. https://doi.org/10.1093/icc/11.3.529

Christiaanse, E., & Markus, M. L. (2002). Business-to-Business Electronic Marketplaces and the Structure of Channel Relationships. *Icis*, 1–21.

Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Platforms and infrastructures in the digital age. *Information Systems Research*, *29*(2), 381–400. https://doi.org/10.1287/isre.2018.0794

Curry, E. (2016). The Big Data Value Chain: Definitions, Concepts and Theoretical Approaches. In *New Horizons for a Data-Driven Economy* (pp. 29–37). Springer, Cham. https://doi.org/10.1201/b19028-5

Deichmann, J., Heineke, K., Reinbacher, T., & Wee, D. (2016). Creating a successful Internet of Things data marketplace. *McKinsey & Company*, *14*(October), 1–8.

Duch-Brown, N., Martens, B., & Mueller-Langer, F. (2017). *The Economics of Ownership, Access and Trade in Digital Data*.

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, *14*(4), 532–550.

European Commission. (2020). *Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - A European Strategy for Data*. https://doi.org/10.1017/CBO9781107415324.004

Flipsen, M. (2019). *Data Collaboratives: Trusted Data Intermediary Business Models*. TU Delft.

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

Galletta, A. (2013). *Mastering the semi-structured interview and beyond: From research design to analysis and publication*. NYU Press.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). Big Data for Big Business? *Cambridge Service Alliance Blog*. http://cambridgeservicealliance.blogspot.co.uk/2014/04/big-data-for-big-business_3.html

Herzfeldt, A., Hausen, M., Briggs, R. O., & Krcmar, H. (2012). Developing a risk management process and risk taxonomy for medium-sized IT solution providers. *ECIS 2012 - Proceedings of the 20th European Conference on Information Systems*.

Hodapp, D., Remane, G., Hanelt, A., & Kolbe, L. M. (2019). Business Models for Internet of Things Platforms: Empirical Development of a Taxonomy and Archetypes. *14th International Conference on Wirtschaftsinformatik, February*, 1783–1797.

Janssen, M., & Verbraeck, A. (2005). *Evaluating the Information Architecture of an Electronic Intermediary*. *15*(1), 35–60. https://doi.org/10.1207/s15327744joce1501

Janssen, M., & Zuiderwijk, A. (2014). Infomediary Business Models for Connecting Open Data Providers and Users. *Social Science Computer Review*, *32*(5), 694–711. https://doi.org/10.1177/0894439314525902

Keller, R., & König, C. (2014). A reference model to support risk identification in cloud networks. *35th International Conference on Information Systems, ICIS 2014*.

Koutroumpis, P., Leiponen, A., & Thomas, L. (2017). The (Unfulfilled) Potential of Data Marketplaces. *ETLA Working Papers*, *2420*(53). http://pub.etla.fi/ETLA-Working-Papers-53.pd

Koutroumpis, P., Leiponen, A., & Thomas, L. D. W. (2020). Markets for data. *Industrial and Corporate Change*, *March*. https://doi.org/10.1093/icc/dtaa002

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage. https://doi.org/10.1002/9781118541555.wbiepc065

Mahoney, J. T. (2004). Economic foundations of strategy. In *Economic Foundations of Strategy*. https://doi.org/10.4135/9781452229850

Mao, W., Zheng, Z., & Wu, F. (2019). Pricing for Revenue Maximization in IoT Data Markets: An Information Design Perspective. *Proceedings - IEEE INFOCOM*, *2019-April*, 1837–1845. https://doi.org/10.1109/INFOCOM.2019.8737571

Martens, B., & Mueller-langer, F. (2018). Access to digital car data and competition in aftersales services markets. *EC JRC Digital Economy Working Paper 2018-06*.

Mayring, P. (2004). Qualitative content analysis. *A Companion to Qualitative Research*, *1*, 159–176.

Mišura, K., & Žagar, M. (2016). Data marketplace for Internet of Things. *Proceedings of 2016 International Conference on Smart Systems and Technologies, SST 2016*, 255–260. https://doi.org/10.1109/SST.2016.7765669

Möller, F., Bauhaus, H., Hoffmann, C., Niess, C., Otto, B., & Isst, F. (2019). Archetypes of Digital Business Models in Logistics Start-Ups. *ECIS 2019 Proceedings*, 1–18.

Muschalle, A., Stahl, F., Löser, A., & Vossen, G. (2012). Pricing approaches for data markets. *International Workshop on Business Intelligence for the Real-Time Enterprise*, 129–144. https://doi.org/10.1007/978-3-642-39872-8_10

Musso, S., Perboli, G., Rosano, M., & Manfredi, A. (2019). A Decentralized Marketplace for M2M Economy for Smart Cities. *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, *i*, 27–30. https://doi.org/10.1109/wetice.2019.00014

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. https://doi.org/10.1057/ejis.2012.26

Nicolaou, A. I., & McKnight, D. H. (2006). Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Information Systems Research*, *17*(4), 332–351. https://doi.org/10.1287/isre.1060.0103

Oberländer, A. M., Lösser, B., & Rau, D. (2019). Taxonomy Research in Information Systems: a Systematic Assessment. *Proceedings of the 27th European Conference on Information Systems (ECIS)*, 0–17. https://aisel.aisnet.org/ecis2019_rp/144

Oberländer, A. M., Röglinger, M., Rosemann, M., & Kees, A. (2018). Conceptualizing business-to-thing interactions–A sociomaterial perspective on the Internet of Things. *European Journal of Information Systems*, *27*(4), 486–502. https://doi.org/10.1080/0960085X.2017.1387714

Ølnes, S., Ubacht, J., & Janssen, M. (2017). Blockchain in government: Benefits and implications of distributed ledger technology for information sharing. *Government Information Quarterly*, *34*(3), 355–364. https://doi.org/10.1016/j.giq.2017.09.007

Osterwalder, A., & Pigneur, Y. (2010). *Business model generation : A handbook for visionaries, game changers, and challengers.*

Park, J. S., Youn, T. Y., Kim, H. Bin, Rhee, K. H., & Shin, S. U. (2018). Smart contract-based review system for an IoT data marketplace. *Sensors (Switzerland)*, *18*(10), 1–16. https://doi.org/10.3390/s18103577

Pavlou, P. A. (2002). Institutional trust in interorganizational exchange relationships: The role of electronic B2B marketplaces. *Journal of Strategic Information Systems*, *11*(3–4), 215–243.

Perera, C., Wakenshaw, S. Y. L., Baarslag, T., Haddadi, H., Bandara, A. K., Mortier, R., Crabtree, A., Ng, I. C. L., McAuley, D., & Crowcroft, J. (2017). Valorising the IoT Databox: creating value for everyone. *Transactions on Emerging Telecommunications Technologies*, *28*(1). https://doi.org/10.1002/ett.3125

Powell, W. W. (1990). Neither Market Nor Hiearchy: Network forms of organization. *Reseach in Organizational Behavior*, *12*(January), 295–336.

Ramirez, E., Brill, J., Ohlhausen, M. K., Wright, J. D., & McSweeny, T. (2014). Data brokers and the need for transparency and accountability. In *Federal Trade Commission*.

Savelyev, A. (2017). Contract law 2.0: ‘Smart’ contracts as the beginning of the end of classic contract law. *Information and Communications Technology Law*, *26*(2), 116–134.

https://doi.org/10.1080/13600834.2017.1301036

Schomm, F., Stahl, F., & Vossen, G. (2013). Marketplaces for Data : An Initial Survey. *ACM SIGMOD*, *42*(1), 15–26.

Smit, B. (2002). Atlas.ti for qualitative data analysis. *Perspectives in Education*, *20*(3), 65–75.

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. *Intereconomics*, *54*(4), 208–216. https://doi.org/10.1007/s10272-019-0826-z

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. *Vietnam Journal of Computer Science*, *3*(3), 137–143. https://doi.org/10.1007/s40595-016-0064-2

Strauss, A., & Corbin, J. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. In *Management Learning* (2nd ed., Vol. 31, Issue 4). Sage. https://doi.org/10.1177/1350507600314007

Susha, I., Janssen, M., & Verhulst, S. (2017). Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*, 2691–2700. https://doi.org/10.24251/hicss.2017.325

Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Because your taxonomy is worth it : Towards a framework for taxonomy evaluation. *Proceedings of the Twenty-Seventh European Conference on Information Systems (ECIS), Stockholm, April*, 0–19.

Täuscher, K., & Laudien, S. (2017). Uncovering the Nature of Platform-Based Business Models: An Empirical Taxonomy. *HICSS 2017 Proceedings*, 5285–5294.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, *43*(2–3), 172–194. https://doi.org/10.1016/j.lrp.2009.07.003

Thomas, L. D. W., & Leiponen, A. (2016). Big data commercialization. *IEEE Engineering Management Review*, *44*(2), 74–90. https://doi.org/10.1109/EMR.2016.2568798

Tiwana, A. (2013). *Platform Ecosystems: Aligning Architecture, Governance and Strategy*. Morgan Kaufmann.

Truong, H. L., Comerio, M., De Paoli, F., Gangadharan, G. R., & Dustdat, S. (2012). Data contracts for cloud-based data marketplaces. *International Journal of Computational Science*, *7*(4), 280–295.

van den Broek, T., & van Veenstra, A. F. (2015). Modes of Governance in Inter-Organizational Data Collaborations. *ECIS*. http://aisel.aisnet.org/ecis2015_cr%0Ahttp://aisel.aisnet.org/ecis2015_cr/188

Williamson, O. E. (1973). Markets and Hierarchies: some elementary considerations. *The American Economic Review*, *63*(2), 316–325.

Williamson, O. E. (1989). Chapter 3 Transaction cost economics. *Handbook of Industrial Organization*, *1*, 135–182. https://doi.org/10.1016/S1573-448X(89)01006-X

Appendices

A. Scientific Paper

A Business Model Taxonomy for Data Marketplaces

Data Trade in Various Trading Structures

Rômy Bergman

Delft University of Technology

ABSTRACT

Companies are increasingly reliant on data to advance their business. Data marketplace owners enable data trade across companies, but little is known about the business models of data marketplaces. Literature mainly focuses on one type of data marketplace with a market orientation and independent ownership, whereas those data marketplaces do not move past the conceptual stage. In this research we investigate the business models of different types of data marketplaces that range from hierarchical to market orientation and private to independent ownership in the B2B automotive industry. The research question states *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?* To answer this question, we create a taxonomy in which we classify six data marketplaces from the B2B automotive industry. Based on business model patterns in our taxonomy we derived four business model archetypes which we link to the data marketplace types. The data marketplaces with private ownership and a hierarchical orientation apply the *aggregating data marketplace* archetype. Data marketplaces with consortium ownership and characteristics from both a hierarchical and market orientation apply the archetypes *aggregating data marketplace with additional brokering service* and *consulting data marketplace*. Data marketplaces with independent ownership and a market orientation apply the *facilitating data marketplace* archetype. To move past the conceptual stage, the data marketplace owners who apply the facilitating data marketplace archetype are advised to offer value adding services such as a customized map service, reviewing data quality or offering personal assistance to their participants. To improve the reliability of the results, researchers may classify additional data marketplace cases in our taxonomy in future research.

Keywords

Data marketplace owner; business model; taxonomy; archetypes; trading structures; automotive industry.

1. INTRODUCTION

An increasing amount of data is being generated by devices. This is a result of the evolvement of the Internet of Things (IoT), which connects devices, systems and people, stimulates the increase in data generation (Cheng, Longo, Cirillo, Bauer & Kovacs, 2015). The emergence of IoT causes a shift from stocks of data into a constant stream of data (Tiwana, 2013). Enterprises become increasingly reliant on data streams as a resource to further advance their businesses (Hartmann, Zaki, Feldmann & Neely, 2014). Companies use internal and external data streams to improve their processes and innovate existing and novel products or services (Agahari, de Reuver & Fiebig, 2019). However, enterprises often only utilize data for primary usage and store it in data silos afterwards (Perera et al., 2017). According to Thomas & Leiponen (2016) the value of data is in its secondary use, when data from organizations is reused externally.

The potential of secondary data use is targeted by data marketplace owners. In theory data marketplaces are multi-sided platforms which allow data sellers, data buyers and third-party service providers to easily trade, store and access data (Mišura & Žagar, 2016). However in reality, companies rarely trade industrial datasets on multilateral data marketplaces and preferably trade data on bilateral basis (Koutroumpis, Leiponen & Thomas, 2017). Several issues that interfere with the advancement of trade via multilateral data marketplaces have been identified such as data security, user privacy (Park, Youn, Kim, Rhee & Shin, 2018; Spiekermann, 2019), data quality preservation (Koutroumpis et al., 2017; Perera et al., 2017), data monetization and revenue optimization (Mao, Zheng, & Wu, 2019; Spiekermann, 2019).

Business models contribute to solving these issues. They provide a framework to understand, analyze and communicate strategic design choices as well as inform the design of information systems (Al-Debei & Avison, 2010). Data marketplace owners apply business models to transform technical ideas into functioning value propositions (Amit & Zott, 2001). A business model demonstrates how companies create, deliver and capture value (Teece, 2010). Literature about business models for data marketplaces is fragmented and is still evolving (Fruhwirth, Rachinger & Prlja, 2020). Various researchers discuss individual components of data marketplaces and propose pricing, quality and privacy mechanisms to improve data marketplaces (Mao et al., 2019; Park et al., 2018; Perera et al., 2017). These proposed mechanisms remain theoretical ideas and are not all implemented in practice (Constantinides, Henfridsson & Parker, 2018). Thus, business model literature about data marketplaces comprehends predominantly theoretical components rather than practically applicable components. Therefore, it remains unclear what business models data marketplace owners actually apply in practice.

Taxonomies are suitable artefacts to provide such insights. As stated by Nickerson, Varshney & Muntermann (2013), taxonomies aid researchers and practitioners in deciding on the uniqueness of existing applications or in pointing out possibilities for new developments. In the emerging research field of data marketplaces, taxonomies are useful to analyze the business models of data marketplaces. However, few taxonomies of business models for data marketplaces exist. Current research contains two taxonomies that have been developed up to by Spiekermann (2019) and Fruhwirth et al. (2020). However, they mainly classified taxonomies that are not past the conceptual phase. Hence, their taxonomies do not provide the insights in business models that data marketplace owners apply in practice. The taxonomies of Spiekermann (2019) and Fruhwirth, Rachinger & Prlja (2020) lack on the following fronts, which asks for an adapted taxonomy:

- Their taxonomies vary in the included business model dimensions. This indicates misalignment or misinterpretation of the dimensions. Spiekermann (2019) includes dimensions such as market positioning and transformation activities, which are excluded by Fruhwirth et al. (2020). On the other hand the taxonomy of Fruhwirth et al. (2020) contains the dimensions time relevancy and payment currency which are not considered by Spiekermann (2019). An explanation why certain dimensions are included or excluded from their taxonomies lacks.
- Their taxonomies are based on cross-industry data marketplaces. This leads to a general interpretation of business model dimensions and characteristics.
- Their taxonomies are based on multilateral data marketplaces that remain conceptual ideas without a viable business model. Spiekermann (2019) and Fruhwirth et al. (2020) do not consider data marketplaces with a hierarchical orientation and private ownership, which have more established business models.

Our objective is to clarify what business models data marketplace owners apply in the business-to-business (B2B) automotive industry. In this research we address the following research question: *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?* We focus on the B2B automotive industry because this industry has established data marketplaces as identified by Martens & Mueller-langer (2018). Through investigating this specific industry, business model components could be identified that data marketplace owners apply successfully in practice. To achieve our objective, we designed a business model taxonomy for data marketplaces and subsequently derived business model archetypes from our taxonomy. We bridge the previously mentioned gaps as follows:

- Our taxonomy includes business model dimensions that we derived from interviews with data marketplace owners. We aligned these dimensions with the dimensions from the taxonomies developed thus far, thereby decreasing variability among taxonomies by Spiekermann (2019) and Fruhwirth et al. (2020).

- We classified data marketplaces from the B2B automotive industry in our taxonomy to be specific in our interpretation of business model dimensions and characteristics.
- We included different types of data marketplaces that vary in their orientation and ownership. The orientation of a data marketplace refers to the coordination of data trade in a hierarchical or market structure. Ownership indicates whether one private company, a number of companies or an independent party owns the data marketplace. In practice, data marketplaces with a hierarchical orientation and private ownership mainly occur. Therefore, we included these data marketplace types, which sets our taxonomy apart from the ones created by Spiekermann (2019) and Fruhwirth et al. (2020).

The remainder of this paper is organized as follows. In section 2 we conceptualize the data marketplace types. We provide theoretical background information about business models and different trading structures in section 3. Next, we explain the research method in section 4. The research results are presented in section 5. Subsequently, we discuss our results and research limitations in section 6. Finally, we conclude the paper in section 7 with an answer to the research question and recommendations for future research.

2. CONCEPTUALIZING DATA MARKETPLACE TYPES

To comprehend what a data marketplace entails, we need to clarify the terms *data* and *marketplace*. *Data* is the core product that is traded at a data marketplace. Data can appear in different forms (raw and aggregated). Stahl et al. (2016) require data marketplaces to contain machine-readable data, such as RDF or XML, which we adopt in our definition of a data marketplace. Platforms where data is traded in textual form, such as Wikipedia, are excluded from our definition. *Marketplaces* are the online or offline infrastructures where marketplace participants exchange goods (Stahl et al., 2016). There are three main functions a marketplace should fulfill (Bakos, 1998):

1. Match buyers and sellers: the buyer’s demand and seller’s supply should be matched by determining the product offerings, searching for buyers and sellers and determining the price.
2. Facilitate transactions: mechanisms for logistics and settlement should lead to the transportation of the sold product and transfer of payment.
3. Provide an institutional infrastructure: markets should have mechanisms to enforce laws, rules and regulations to coordinate transactions.

With the previous notions of *data* and *marketplaces* we create the following definition of a data marketplace: a *data marketplace* <u>matches</u> buyers and sellers, <u>facilitates</u> transactions and <u>provides</u> an institutional infrastructure to trade <u>machine-readable data</u>.

In literature, researchers characterize data marketplaces by the participants who are active on the platform. Four key players are mentioned frequently. These are the data marketplace owner, data sellers, data buyers and third party service providers (Fruhwirth et al., 2020; Koutroumpis et al., 2017; Muschalle et al., 2012; Spiekermann, 2019). Spiekermann (2019) describes the relationships between the four key players. The data marketplace owner hosts the data on the platform. The data is made available by the data seller, who owns the data. Data sellers may be commercial or non-commercial parties. Data is sold to data buyers who are consumers or businesses. Third party service providers leverage datasets and add value to the data. They retrieve data from the data marketplace and upload a transformed dataset.

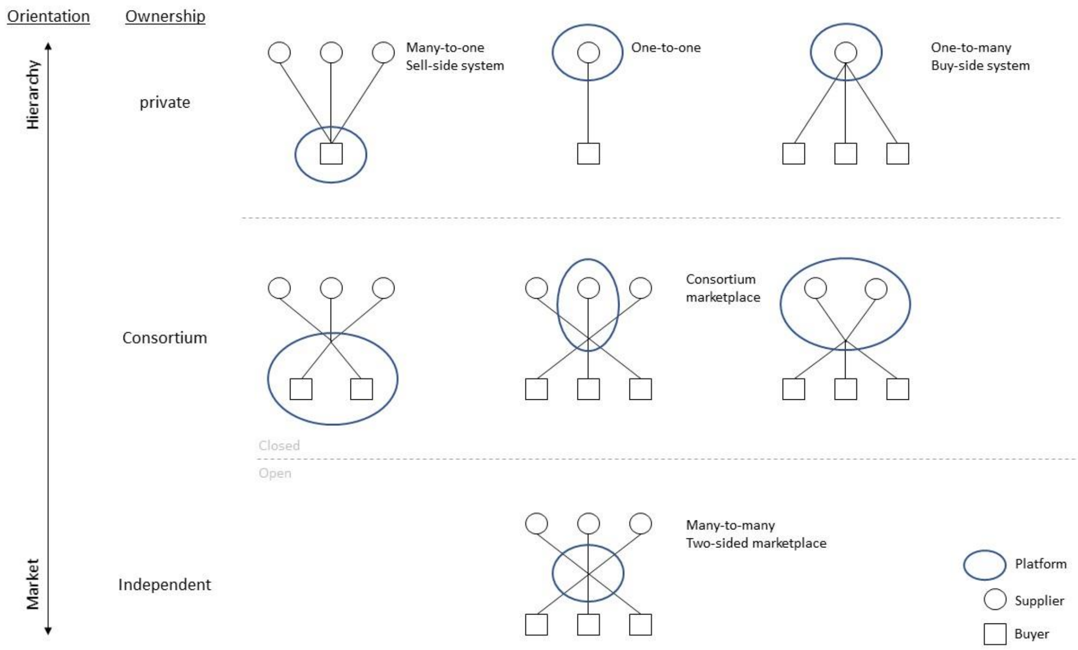

Figure 1: Data marketplace types adapted from Koutroumpis et al. (2017) and Stahl et al (2016)

According to Spiekermann (2019), data marketplace participants trade data on multilateral basis. As explained in section 1, multilateral data marketplaces remain theoretical. In reality data marketplace participants rather trade data on bilateral basis. To research business models of data marketplaces that go beyond theoretical concepts, we include business models of data marketplaces that occur in practice. This indicates that we need to consider different types of data marketplaces.

Stahl et al. (2016) propose a framework that enables classification of data marketplaces in different types. They make use of two determinants: orientation and ownership. Orientation refers to whether the data marketplace owner coordinates data trade in a hierarchical or market trading structure. In data marketplaces with a hierarchical orientation, the owner determines the data price and what buyers and sellers are allowed on the platform. In the data marketplace with a market orientation, prices are determined by the buyers and sellers depending on competitive offerings. Ownership indicates whether one private company, a number of companies or an independent party owns the data marketplace. Koutroumpis et al. (2017) maintain a similar classification in which they sort data marketplaces based on their matching mechanism. They distinguish between four types of data marketplaces; one-to-one, one-to-many, many-to-one and many-to-many data marketplaces. First, one-to-one data marketplaces are bilateral marketplaces where two parties are directly connected. One seller will trade with one buyer. Second, at one-to-many data marketplaces there is one seller who trades with many buyers for the same data. In this case, standardized terms of exchange through APIs are maintained, because it is too costly to negotiate data individually. Third, many-to-one data marketplaces allow multiple sellers and one buyer. The sellers usually make their data available to one service provider and receive a service in return for free, as practiced on social media platforms. Fourth, many-to-many data marketplaces are multilateral marketplaces where many sellers and buyers trade data. There is often no specific ownership over the data, but transactions to acquire data are facilitated.

We combine the classifications of Koutroumpis et al. (2017) and Stahl et al. (2016) in figure 1. It forms a spectrum in which different types of data marketplaces can be classified, depending on their orientation and ownership. This allows us to research data marketplaces that range from hierarchical to market orientation and private to independent ownership. We extend our previous definition of a data marketplace to: a *data marketplace* has a hierarchical or market <u>orientation</u> and private, consortium or independent <u>ownership</u> and <u>matches</u> buyers and sellers, <u>facilitates</u> transactions and <u>provides</u> an institutional infrastructure to trade <u>machine-readable data</u>.

3. THEORETICAL BACKGROUND

3.1 Business Models

To create a business model taxonomy for data marketplaces, consensus about the representation of a business model should be established. Chesbrough & Rosenbloom (2002) describe business models as frameworks that convert technological input into economic output. Hence, business models have the ability to transform technical potential into economic value. Teece (2010) defines business models as the design in which value is delivered to the customer. Amit & Zott (2001) explain that business models visualize the design of "transaction content, structure and governance", to create value from various sources and discover new business opportunities. Despite the variety in business model definitions, all business model descriptions include component-based perspectives (Hartmann et al., 2014). Therefore, we represent business models based on the business model components.

The business model canvas by Osterwalder & Pigneur (2010) is chosen as leading framework in this paper to identify the main business model components, because it combines all aspects that are identified by Amit & Zott (2001), Chesbrough & Rosenbloom (2002) and Teece (2010). We classify the business model components under the main components; value creation, value delivery and value capture (Teece, 2010). First, value creation is the process of making something that brings worth to the customer. We assign the components value proposition, customer segment and customer relationships to the main component value creation. Second, value delivery is about the asset arriving at the customer. The channels, key resources, key activities and key partners contribute to value delivery of data marketplaces. Third, when data marketplace owners capture value, they monetize the created and delivered value. We assign the components revenue streams, pricing model and cost model to value capture. These business model components are further specified for data marketplaces in business model dimensions and characteristics in our taxonomy.

3.2 Trading Structures

Data marketplaces differ in their orientation. As explained in section 2, we define data marketplace types with a hierarchical and market orientation. Because our types are based on the classification of Stahl et al. (2016) and Koutroumpis et al. (2017), we focus on the hierarchical and market orientation in our definition of data marketplace types. Williamson (1973; 1989) established the market-hierarchy continuum to explain factors that cause a shift from a market to a hierarchical structure. We use those factors to characterize the hierarchical and market orientation of data marketplaces.

First, bounded rationality refers to the inability of humans to "receive, store, retrieve and process information without error" (Williamson, 1973, p. 107). Although humans try to act rational, limited information makes them reach a satisfactory solution instead of an optimal solution. In a hierarchical structure, bounded rationality poses less of a problem than in a market structure. The clear boundaries in departments, lines of authority and reporting mechanisms internalize transactions in a hierarchical structure. This enables these organizations to "write contracts that cover all possible contingencies" (Powell, 1990, p. 297). In a market structure, boundaries are less clear. Bounded rationality in organizations with a market structure make them prone to market failure.

Second, opportunism entails the aim of actors to maximize their personal gain (Williamson, 1973). Actors can go as far as deceit to achieve their goal. Powell (1990) explains that actors aim to minimize their costs at in organizations with a market structure. Production and exchange at these organizations are determined by price competition. When the price of a product does not satisfy the needs of an actor in a market structure, he has the flexibility to move to another seller who does meet his requirements. Powell (1990) notes that interactions in a market structure do not "establish strong bonds of altruistic attachments" (p. 302). This leads to quick and efficient interactions with a lack of strong relationships. On the other hand, in hierarchical structures actors practice authority by imposing rules and sanctions to regulate opportunistic behavior (Williamson, 1989). According to Powell (1990), actors communicate based on routines with people they are familiar with. As such, people who "know one another, have a history of previous interactions and possess a good deal of firm-specific knowledge" trade in hierarchical structures (Powell, 1990, p. 302). Authoritative relations and personal identification pose less room for opportunistic behavior in a hierarchical structure. According to Williamson (1973), opportunism does not lead to the success of one structure over the other. It indicates what actors will more likely trade in what structure. People who aim for maximum personal gain are attracted to a market structure and people who seek routine are attracted to a hierarchical structure.

Third, uncertainty in a market structure influences economic behavior (Williamson, 1973). Examples of uncertain factors at the time of exchange in a market structure are future price, demand/supply ratio and price/quality estimation. This may lead to non-optimal transactions. Powell (1990) explains that there is more control over coordinating supply and demand in a hierarchical structure, stating “the visible hand of management supplants the invisible hand of the market in coordinating supply and demand” (p. 303). He explains that managerial teams in a hierarchical structure have the ability to coordinate high volume and speed operations. The vertical integration of organizations with a hierarhical structure enables them to well-coordinate mass production and distribution. This mitigates uncertainties.

Fourth, small numbers are unfavorable for a market structure (Williamson, 1973). A high number of buyers and sellers creates a competitive environment in market structures, which stimulates price reduction. Small numbers cause an organization with a market structure to shrink or vanish altogether.

In addition to the human and transactional factors, Williamson (1989) introduces asset specificity. This entails the extent to which an asset can be used for multiple purposes. Williamson recognizes five different forms of asset specificity. These are site specificity, physical asset specificity, human asset specificity, dedicated assets and brand name capital. Powell (1990) elaborates that “transaction-specific investments – of money, time and energy that cannot be easily transferred – are more likely to take place within hierarchically organized firms” (p. 297). Thus, as assets become more knowledge specific, they are likely traded in organizations with a hierarchical structure.

We use the previous notions of asset specificity and opportunism to define organizations with a hierarchical or market structure. Asset specificity is a relevant concept for the value proposition of a data marketplace. *Asset specificity* concerns the goods that are traded in a hierarchical or market structure. Organizations with a hierarchical structure likely trade in asset specific. Organizations with a market structure trade in less asset specific goods. Opportunism is a relevant factor to describe the customers who are attracted to a data marketplace. *Opportunism* concerns the people who are attracted to organizations with a hierarchical or market structure. Organizations with a hierarchical structure trade on authoritative basis between actors who are familiar with each other. Organizations with a market structure trade on competitive basis between actors who aim for the maximum individual gain. This leads to the following definitions of organizations with a hierarchical and market structure and continuum in which data marketplaces can be classified as visualized in figure 2.

i. *Organizations with a hierarchical structure* trade in asset specific goods on authoritative basis between actors who are familiar with each other
ii. *Organizations with a market structure* trade in less asset-specific goods on competitive basis between actors who aim for the maximum individual gain

Figure 2: Data marketplaces in the market – hierarchy continuum

Based on the presence of two distinctive orientations of data marketplaces in practice and literature, data marketplaces with a hierarchical orientation and market orientation are included in this paper. In practice, data marketplaces with a hierarchical orientation exist. Attempts are made to launch multilateral data marketplaces with a market orientation in practice, but these initiatives do not yet succeed (Koutroumpis et al., 2017). In literature, scientists focus on data marketplaces with a market orientation to advance the development of these data marketplace types (Fruhwirth et al., 2020; Schomm et al., 2013; Spiekermann, 2019). Our research contribution is to identify the business models of the different types of data marketplaces. We will clarify our findings in the discussion with respect to the market – hierarchy continuum of data marketplaces and suggest future research directions.

4. METHOD

To identify the business models that data marketplace owners apply, we performed a qualitative research study. A business model taxonomy is designed in which we classified business models of various data marketplace types. The taxonomy development approach by Nickerson et al. (2013) offers a systematic way to create a taxonomy and is widely accepted in the field of information systems (Szopinski, Schoormann & Kundisch, 2019). Therefore, we followed their iterative approach and combined inductive and deductive research (see figure 3). In the empirical-to-conceptual approach, concepts are induced from existing objects. In the conceptual-to-empirical approach, concepts are deduced from literature. The combination of both approaches led to the design of our taxonomy.

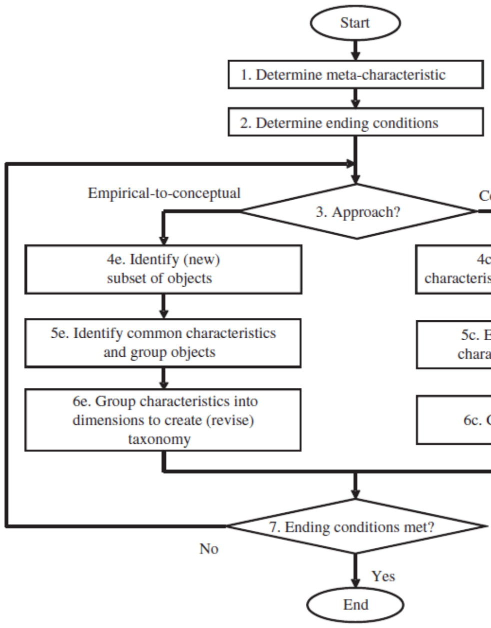

Figure 3: Taxonomy development approach by Nickerson et al. (2013)

Before we induced or deduced concepts for our taxonomy, meta-characteristics were determined. The choice in meta-characteristics should be based on the purpose of the taxonomy (Nickerson et al., 2013). The purpose of our taxonomy is to identify the business model dimensions and characteristics that data marketplace owners in the B2B automotive industry apply. Thus, the meta-characteristics of our taxonomy need to represent the main business model components. Teece (2010) describes value creation, delivery and capture as the main business model mechanisms. These mechanisms are the chosen meta-characteristics for our taxonomy. In addition, we adopted the objective and subjective ending conditions of Nickerson et al. (2013), as done by Fruhwirth et al. (2020) and Möller et al. (2019). The ending conditions are listed in table 1.

Table 1: Taxonomy iterations

Ending Conditions		Iterations							
		1	2	3	4	5	6	7	8
OE1	All objects or a representative sample of objects have been examined							x	x
OE2	No object was merged with a similar object or split into multiple objects in the last iteration	x	x	x	x	x	x	x	x
OE3	At least one object is classified under every characteristic of every dimension					x	x	x	x
OE4	No new dimensions or characteristics were added in the last iteration							x	x
OE5	No dimensions or characteristics were merged or split in the last iteration	x	x	x		x	x	x	x

OE6	Every dimension is unique and not repeated (i.e., there is no dimension duplication)						x	x	x
OE7	Every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension)						x	x	x
SE1	Concise: the taxonomy is meaningful without being overwhelming								x
SE2	Robust: the dimensions and characteristics suffice to differentiate objects								x
SE3	Comprehensive: all objects can be classified								x
SE4	Extendible: new dimensions and characteristics can be added								x
SE5	Explanatory: the dimensions and characteristics explain an object								x

In iteration 1 we deduced general business model components. This comprises the business model components from the Business Model Canvas by Osterwalder & Pigneur (2010). The business model components are further specified in our taxonomy for data marketplaces.

In iteration 2 business model dimensions were induced from interviews with data marketplace owners. These dimensions were induced using the Grounded Theory Method. Grounded theory is constructed through inductive reasoning, starting with information gathered from interviews, reports and other data materials. This explorative start is advised by Nickerson et al. (2013) in areas where little data about the research domain is available. We conducted seven interviews with data marketplace owners to learn about their business models (see table 2). We aimed for variance in the data marketplace types that the interviewees represent. The types range from hierarchical to market orientation and private to independent ownership, which is introduced in section 2. At minimum, each marketplace type must be represented by at least one interviewee. In addition, we maintained the following selection criteria:

- The interviewees are available for an interview and speak English
- The interviewees work at data marketplaces that trade in automotive data.
- The interviewees have business model knowledge. We judged this based on the job title of the approached interviewees. Interviewees who occupy a position related to business development, are expected to have in-depth knowledge about the business model.
- The interviewees have over 5 years of work experience in business development or consultancy. Because many data marketplaces are newly founded in the last 5 years, we could not expect the interviewees to have over 5 years of work experience at the respective data marketplace. Therefore, we looked at their work experience previous to their current job.

Table 2: Grounded Theory interview respondents

Code	Type	Job title	Other relevant experience
DM1	Mixed hierarchy market, consortium	Business development	Previously worked as a marketing and business development consultant for 4 years
DM2	Market, independent	Product owner	Over 5 years of experience as a data scientist and business consultant at various multinationals
DM3	Market, independent	Unknown	7 years of experience in advising ministries about traffic and mobility data
DM4	Market, independent	Business Development	Over 5 years of experience as a freelance consultant
DM5	Market, independent	Innovation Manager Smart Mobility	Previously worked as a consultant for national agencies and has over 5 years of experience working on smart mobility projects
DM6	Market, private	Director Business Development	Over 10 years of experience at various IT service providers as sales manager
DM7	Hierarchical, private	Head of Enterprise Business Development	Over 8 years of experience in corporate development at a multinational

During intensive interviewing we encouraged, listened to and learned from the interviewees. This method is suggested by Charmaz (2006) to construct grounded theory. Charmaz (2006) developed a theory building process that starts with open ended interview questions. Based initial interview questions suggested by Charmaz (2006, p. 30-31), four initial questions were formulated to start the conversation. We asked the interviewees to describe the trends and challenges that their data marketplaces face, how their business model responds to those trends and challenges and what the difference is between their data marketplace and a competitor. Based on the responses to our initial questions, we formulated intermediate questions as the interviews proceeded. This led to open-ended and in-depth exploration of the interviewees' area of expertise. Qualitative coding was applied to construct the interview data into categories that represent business model concepts. This is described by Charmaz as "the process of defining what the data are about" (2006, p. 43). During this step, the process of data collection is linked to theory building. Charmaz (2006) identifies two main coding phases that we followed: initial coding and focused coding.

First, we named data segments line-by-line during initial coding. For example, one interviewee described their data processing activities saying *"Data marketplaces often need to do data aggregation before giving data to the user. We use data cataloguing for this process. There you can do data tagging and data cleansing"* (DM2). We assigned the following codes to these lines: *aggregate data*, *catalogue data*, *tag data* and *clean data*.

Second, we searched for the most frequent or significant codes during focused coding. We separated, sorted and synthesized the line-by-line codes into categories for a text segment. Focused coding is the process during which the researcher starts to recognize relationships and patterns between categories (Charmaz, 2006). For example, based on the initial codes of the interviewee describing the data processing activities, we selected *aggregate data* as the most significant code for this piece of text.

After this process was completed for all interviews separately, we performed a second round of focused coding to construct categories that apply to all interviews. This second step of focused coding was required, because interviewees used different wordings to describe similar processes. For example, we created the focused codes *searching databases*, *aggregate data* and *harmonize and synchronize data* that stem from different interviews. These codes all refer to the data processing activities that a data marketplace owner performs. Hence, in the second round of focused coding, the overarching category *data processing activities* emerged. Likewise, six more categories emerged: *data regulation*, *customers*, *platform infrastructure*, *revenue model*, *data quality* and *other* (see table 3).

Table 3: Constructed categories

Main categories	Focused codes
Data regulation	Design quality standards / smart contract / comply with GDPR / delegated data regulation / preserve data privacy / comply with EU law / setting legal framework is challenging / terms and conditions / privacy is a challenge / privacy regulation / privacy challenge / check privacy regulation / terms and conditions determine data usage / Privacy disables open data publication / use of data is a license / non-cooperation of OEMs / customers restrict data usage / data ownership
Customers	Users / industry domain / attracting a specific customer segment / large target group / customer segment / maintain customer segments / direct OEM relationship
Platform infrastructure	Decentral data control / open governance / centralized or decentralized approach / open protocol / open platform infrastructure / decentral infrastructure / decision making at consumer / challenge to regulate IT integration
Data processing activities	Searching databases / saving time / overview in catalog / perform additional activities / provide corporate and open data / aggregate data / perform activities for all needs / advertise meta data / mixed functionalities / preserve data privacy / national access point / offer broker services / regulate data availability / broker of data / harmonize and synchronize data / setting legal framework is challenging / data processing activities / acting as traditional marketplace / extracting value from data is a challenge / acting as a consent management hub / key activities / key processes / perform several activities / advise OEM in data supply / market leader for development and research / develop own data products / perform activities / enable analysis of car data / generating insights from sensor data is difficult / collect data which is needed / differentiate added value / expand product offering / value chain depends on layer / multiple suppliers cause more activities / production is partially standardized and partially customized / try to standardize terms and conditions / high service quality / added value of aggregated product / performance of data processing activities

Revenue model	Valorizing data is difficult / no active role in data pricing / price discovery / explaining the crypto currency is a challenge / pricing / online product prices / data licensing / Licensing disables open data publication
Data quality	No check of data quality / cooperate to create quality standards / cooperate to automatically improve data
Other	Data marketplace concept / small company size / explain data marketplace concept / role MDM / create fit between customer and marketplace / data marketplace type / fit between governance and client / increase in data generation / more hardware in vehicles / evolve mobility definition / cooperate to develop solution / More data collection because of partnerships

In iteration 3, dimensions were deduced from the existing taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). We did not start with deducing concepts, because two taxonomies do not provide a sufficient amount of data to base our taxonomy on. By gathering data from interviews first, we could learn about relevant business model dimensions for data marketplace owners in practice and supplement these with dimensions from theory. We aligned the induced and deduced dimensions to create our preliminary taxonomy. Our preliminary taxonomy was still conceptual, which means that existing data marketplace cases were not yet classified in our taxonomy.

In iterations 4-6 we refined our preliminary taxonomy with induced business model characteristics from a selection of existing data marketplaces. Through theoretical replication, based on orientation and ownership, we searched for a variety of cases that are spread over three cells of data marketplace types. Additionally, we selected two cases per cell for theoretical sampling. This results in six data marketplace cases that we analyzed from three data marketplace types. These are (i) data marketplaces with a hierarchical orientation and private ownership, (ii) data marketplaces with a mixed hierarchy and market orientation and consortium ownership and (iii) data marketplaces with a market orientation and independent ownership. We limited our analysis to a number of six data marketplaces to perform in-depth case analyses and create more specific business model insights than currently available in the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). Their taxonomies are based on 16-20 data marketplaces. This leads to general a general interpretation of the characteristics in their taxonomies. To find cases from the data marketplace types in the B2B automotive industry, we performed online desk research. The following selection criteria were applied:

- The data marketplaces fit our definition of a data marketplace from section 2: a data marketplace matches buyers and sellers, facilitates transactions and provides an institutional infrastructure to trade machine-readable data
- The data marketplaces trade in automotive data
- The data marketplaces are B2B
- Case documentation is in English
- Past the conceptual phase

Due to the low number of data marketplaces in practice, there were few cases to choose from. We included TomTom, INRIX, HERE, Caruso, IOTA and Ocean Protocol in our case sample (see table 4). It must be noted that IOTA and Ocean Protocol are data marketplaces in the conceptual stage and trade data across industries. We could not find data marketplaces with a market orientation and independent ownership that trade in automotive data and are past the conceptual stage. Ocean Protocol is included in our case sample because the data marketplace is in the beta phase, almost ready for final release. IOTA is included due to their high number of 70 signed up participants.

Through content analysis, we induced business model characteristics from case documents. The webpages, whitepapers and terms of use documents contain the most important information from the point of view of the data marketplace owner. We analyzed these sources first to get an impression of the vision and activities of data marketplace owners. Additionally, articles from external sources were consulted. Forbes is selected as external source, because it is a renowned company, focusing on business, investing, technology, entrepreneurship, leadership and lifestyle. If we required additional information after analyzing these sources, news releases of the cases were included until we reached saturation. An overview of the selected sources is presented in table 5.

Table 4: Case descriptions

Data marketplace	Description
TomTom (TT) Hierarchical, private Founded: 1991	TomTom is a privately owned company that uses location technology to sell mapped data. They trade data in a hierarchically oriented, bilateral market. TomTom is well known for their sale of navigation boxes to end consumers. However in this study, the focus is on the B2B segment of TomTom that concerns data trade between TomTom and their commercial buyers and seller.
INRIX (IN) Hierarchical, private Founded: 2005	INRIX is also a privately owned company and applies location analytics to make road transportation more intelligent. INRIX trades data bilaterally with their commercial data sellers and buyers and serves public organizations. In addition to trading data, INRIX performs research on subjects such as road congestion, commuting time and vehicle carbon emission. Their research branch is out of scope of this thesis, because the reports are in textual format and are not considered machine-readable data.
HERE (HE) Hierarchical/market, consortium Founded: 2015	HERE is formerly known as Navteq and was owned by Nokia. In 2015 the company was acquired by a consortium. Audi, BMW and Daimler are the main shareholders. HERE applies location technology to improve connected driving experiences. The HERE data marketplace has open access for any data seller, data buyer and third-party service provider to exchange data.
Caruso (CR) Hierarchical/market, consortium Founded: 2017	Caruso is founded by TecAlliance, a provider of vehicle data in the automotive industry. Besides TecAlliance, companies such as Bosch and Continental are shareholders of Caruso. The data marketplace is closed and only the consortium members and partners are allowed to trade at the data marketplace.
IOTA Market, independent Founded: 2017	IOTA is founded by the non-profit IOTA Foundation. This data marketplace focuses on the IoT market with the goal to enable secure data transactions between data sellers and buyers. The IOTA data marketplace has open access that allows many participants to trade data. It is currently in the proof of concept phase.
Ocean Protocol (OP) Market, independent Founded: 2017	Ocean Protocol is a non-profit organization based in Singapore. Their data marketplace has open access to create an environment in which many data sellers and buyers can exploit data. The data marketplace is currently in its beta stage and is planned for a new release in Q3 of 2020. Ocean Protocol is particularly focused on AI. With high volumes of data and trained algorithms they aim to advance AI development.

Table 5: Case sources

Sources	TomTom	INRIX	HERE	Caruso	IOTA	Ocean Protocol
Website	Main website develop portal	Main website	Main website Develop portal	Main website	Marketplace platform	Main website
Terms and conditions	Buyer Supplier	Site terms	Service terms	Privacy	Privacy	Privacy
Whitepaper	Product Annual report	Product	Product	Slides Live presentation	Technical	Technical Business Marketplace
External	Forbes articles	Forbes articles Harvard Business Review	Forbes articles	Automat report	Forbes articles	Forbes articles
Total	*9*	*8*	*7*	*5*	*5*	*7*

In iterations 7-8 we revised the dimensions and characteristics in our taxonomy. The objective ending conditions were revised throughout the entire taxonomy development process. In iteration 7, all data marketplaces were classified and all objective ending conditions were met. To assess the subjective ending conditions, we conducted semi-structured interviews with Spiekermann (2019) and Fruhwirth et al. (2020) who we consider experts in the development of business model taxonomies for data marketplaces. We prefer the semi-structured interview method over the structured or unstructured interview methods, because semi-structured interviews allow new topics to be explored while still evaluating predefined conditions (Galletta, 2013). Similarly, Keller & König (2014) applied semi-structured interviews to evaluate their taxonomy which proved as a suitable method to test the ending conditions. Although our taxonomy is intended to be used by researchers and practitioners, we only evaluated our taxonomy with researchers due to time constraints. We recommend taxonomy evaluation with practitioners for future research. Based on the feedback from the experts we adapted our taxonomy and met all subjective ending conditions in the eight iteration.

5. RESULTS

5.1 Preliminary Business Model Dimensions

From interviews with data marketplace owners, we learned about the business models of data marketplaces. We derived five dimensions that data marketplace owners apply to create, deliver and capture value; contract, platform infrastructure, data processing activities, revenue streams and data pricing mechanism. These are the first set of preliminary dimensions that we included in our taxonomy.

Regulating data trade in a contract Most interviewees acknowledged data regulation as a challenge, because rules are not always clearly defined and leave room for interpretation. The interviewees explained that they incorporate rules to preserve data privacy and adhere to regulation. One data marketplace owner explained that they protect data privacy by anonymizing data that is stored: *"stored data is anonymized in such a way that one cannot derive what car would drive to what address and what individual the information belongs to"* (DM6). Another data marketplace owner explained that the data they trade, complies with GDPR. They do not store privacy sensitive information about their users: *"all data items that are put in the data marketplace comply with GDPR. We do not store any user data"* (DM2). In addition to storing data in a secure manner, data marketplace owners incorporate data regulation into their contracts. For example, data marketplace owners agree with their data sellers on the data that will be traded at the data marketplace. Consent is given to trade data, as an interviewee explained: *"we function as a consent management hub. We facilitate communication between a newly developed application and an Original Equipment Manufacturer to give consent to use parameters of a car"* (DM6). The rules to trade data at the data marketplaces are clarified in their terms and conditions: *"everyone comes here to do business and we have clear terms and conditions that say how to trade data"* (DM4). From these statements we derived that although data regulation is identified as a challenge, data marketplace owners learned how to interpret regulation and incorporate rules into their contracts. Besides applying contracts to adhere to data regulation, data marketplace owners use contracts to build customer relationships. This is explained in the following paragraph.

Maintaining trust and power relationships in a contract A second observation that we made is that data marketplace owners use their contracts to establish trust and power relationships. A data marketplace owner (DM1) explained that they maintain strict privacy rules which they communicate to their customers. If data buyers request data that violates privacy regulation, the data marketplace owner denies this request. Another data marketplace owner explained that they are transparent towards data buyers about the usage terms in their contract: *"the usage terms define whether the data seller imposes copyrights for certain regions"* (DM7). Both these examples constitute institutional trust that the data marketplace owner creates by means of transparent usage conditions in their contracts. Furthermore, the playing field in the automotive industry is dominated by power relationships (Martens & Mueller-langer, 2018). According to Martens & Mueller-langer (2018), OEMs who have monopoly power are not eager to participate at a data marketplace because they fear to lose exclusive access to their data. To stimulate participation of OEMs and preserve power at the data seller, some data marketplace owners incorporate usage conditions into their negotiated contracts with OEMs. One of the interviewees explained that their data sellers can demand usage conditions in their contracts that restrict certain participants from accessing the data, stating *"certain sources can set specific conditions for the use of data for applications. Our data suppliers can exclude data usage for specific applications, regions, type of vehicle or end user"* (DM7). The OEMs can exert power through negotiated contract conditions. Another interviewee thinks that without government control, powerful stakeholders in the automotive industry may never share their data at an open data marketplace. He implied that companies can be forced to make their data available at data marketplaces, which he called national data access points, via new regulation: *"With newer regulations, data needs to be made available to national access points in Europe. [...] We are discussing how to better reach international organizations with other European member states to make sure that data provision to the access points is known and is considered in regards to the EU law"* (DM3). However, such regulations do not exist yet. Overall, contracts are used by data marketplace owners to build relationships with their customers. Therefore, we include *contract* as a dimension in our taxonomy, part of the meta-characteristic *value creation*.

Storing data in a centralized or decentralized platform infrastructure Data marketplace owners use their platform infrastructure as a resource to store data. In centralized platforms, data control shifts towards the data marketplace owner who manages the storage location. The data marketplace owners use the centralized platform infrastructure as a resource for data analysis. In decentralized platform infrastructures, the data seller maintains data control. One interviewee said the following about their decentralized platform infrastructure: *"we do not know where the data is. We only are the protocol in between that helps transactions happen"* (DM4). They deploy a decentralized platform infrastructure without a central point of control. They do not store or process data. A decentralized platform infrastructure increases data sovereignty for the data seller and buyer but complicates data storage and analysis for the data marketplace owner. We include *platform infrastructure* as a dimension in our taxonomy, part of the meta-characteristic value delivery.

Performing data processing activities to transform data We identified six data processing activities that data marketplace owners may perform to transform the data traded on their platform. These are data collection, data standardization, data cleansing, data storage and data analysis. It must be noted that not all interviewees expressed to perform all these activities. Some only performed data collection and delivery and let their customers standardize, clean and analyze data. First, during data collection, data marketplace participants agree on the data that will be traded and processed. One interviewee explained that during data collection, terms and conditions are agreed on: *"The data that we buy is always accompanied with terms and conditions. We try to standardize those as much as possible. However, there is a difference in terms and conditions among suppliers. So, the moment a customer buys one of our products, the agreement states under what conditions the data may or may not be used"* (DM7). The conditions that data marketplace owners, sellers and buyers agree upon during data collection, influence what data processing activities may be performed. Second, data is standardized to enable easy exchange of data. One interviewee explained that they format data from various sources into one type: *"...we facilitate IT integration. We enable standardization of the data in such a way that it results in one common language to easily deliver data to consumers"* (DM6). Third, data is cleaned. During data cleansing, data marketplace owners check the data consistency and verify the data content. One data marketplace owner explained that they clean data in collaboration with their customers: *"A challenge for most of our customers who make use of digital maps is to ensure that the provided data is correct. We collaborate with our customers to detect data that is incorrect and automatically improves this in the system. This brings us the advantage to improve the digital map without manual interaction"* (DM7). Other data marketplace owners do not clean the data and let their participants do quality checks. One interviewee explained that their data buyers rate the quality of the datasets that are sold at their data marketplace: *"We designed quality standards. With those standards, data is rated"* (DM2). Data buyers can refer to the quality ratings to estimate the data quality they can expect from a data set. Fourth, data is stored at a secure location that is scalable. Storage facilities have to cope with the volume, velocity and variety of data. As explained before, data can be stored in a centralized or decentralized platform infrastructure. Fifth, during data analysis, data marketplace owners can aggregate and analyze the datasets to extract new insights. One interviewee explained that they aggregate and analyze the data of their data sellers: *"we have many data suppliers. We process data, remove mistakes from the data, link data together and sell this as an aggregated product"* (DM7). Another interviewee explained that they provide tools for their customers who use the tools and perform data analysis themselves: *"Business* *users get visualization tools to derive the data into a graph tool. They can analyze the data and get the visuals"* (DM2). Sixth, data is distributed to the participants. One data marketplace owner explained that they only collect and distribute data: *"our data marketplace has two main functionalities. One is to show the meta-data of available datasets. The other is the brokerage functionality. That is to get data from a data provider and distribute this to all data users who need to subscribe to a data publication. So it's a data delivery and brokerage service"* (DM3). Overall, we noticed that there is a difference in data marketplace owners who perform all data processing activities or a limited number of data processing activities. This influences the data that is delivered to the customer. We include the dimension *data processing activities* in our taxonomy, part of the meta-characteristic value delivery.

Revenue streams to generate income Data marketplace owners generate income from their revenue streams. A data marketplace owner can receive numerous revenue streams. For example, they may charge customers for the usage of their marketplace and customers can be charged for the data that is transferred. Five revenue models are generally applied (Muschalle et al., 2012): free, usage-based process, package pricing, flat fee tariff and freemium. Data marketplace owners can combine these models to generate income. One interviewee explained that they combine the freemium model with a usage based model (DM1). Developers can get data up to a limit of 250.000 transactions per month for free. When they exceed this limit, customers pay a price per data volume that is transferred. We include *revenue streams* as a dimension in our taxonomy, part of the meta-characteristic value capture.

Monetizing data with fixed or dynamic pricing mechanisms Data monetization is perceived as a challenge by data marketplace owners. As one interviewee explained: *"people do not know how to value data. This is a problem. You cannot have a marketplace where you do not know the value of what you are selling"* (DM4). Overall, data marketplace owners could apply two types of pricing mechanisms. These are fixed pricing and dynamic pricing. When fixed pricing mechanisms are applied, the data price is predefined and static. As explained by an interviewee, they trade data based on fixed prices: *"the data price is predefined and the total price is determined based on how much data the data seller consumed"* (DM6). Data marketplace owners who apply dynamic pricing mechanisms aim for data sellers to become price takers. In order for dynamic pricing to succeed, an interviewee explained: "*There is a need for price discoveries and mechanisms that calculate liquidity based on the market and come up with the price. This is still very abstract"* (DM4). Although dynamic pricing is what this data marketplace owner strives for, they apply fixed pricing models in practice: *"Fixed pricing is the easiest play in the book. Come up with a number, and see if people are interested or drop the price. But this is definitely not the solution, because it's not in people's normal workflow to go and put a price on data. Nobody knows how to do this"* (DM4). We include the dimension *data pricing mechanisms* in our taxonomy, part of the meta-characteristic value capture.

We aligned our preliminary dimensions with deduced dimensions from the taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020). This results in a number of preliminary dimensions (see table 6). In section 5.2, we present our final taxonomy in which existing data marketplaces from the B2B automotive industry are classified.

Table 6: Preliminary business model dimensions

Preliminary dimension	**Description**
Domain	the market in which the data marketplace is active
Participants	the data sellers and buyers who are matched at the data marketplace
Data source	the governmental, social media, self-generated or community source where the data on the data marketplace is collected from
Data output	the aggregated or standardized data offering
Data quality	the user reviews or reviews by the data marketplace owner to guarantee data quality of the traded data
Privacy	the anonymization or encryption of data to protect data privacy
Contract	the negotiated or standardized agreements that regulate data trade
Platform access	the open or closed platform access for the customers
Platform infrastructure	the centralized or decentralized storage location of data
Data processing activities	the performance of all or a limited amount of activities by the data marketplace owner to increase the value of the traded data

Revenue streams	the manner in which the data marketplace owner generates income by applying usage based, package pricing, flat fee tariff or freemium models
Data pricing mechanism	the fixed (set by data marketplace owner, sellers or buyers) or dynamic (auction, negotiation, real-time market) pricing mechanism of the data output
Payment currency	the fiat or cryptocurrency in which payments are transferred

5.2 Business Model Taxonomy

Our final taxonomy contains thirteen business model dimensions (see table 7). We removed the preliminary dimension *data source* from our final taxonomy because the characteristics of this dimension overlapped with the characteristics of the dimension *participants*. A newly added dimension is *data service*. The data service that data marketplace owners perform shapes their value proposition and is an integral part in their business model. We specified the business model characteristics of TomTom (TT), INRIX (IN), HERE (HE), Caruso (CR), IOTA and Ocean Protocol (OP) in our taxonomy. According to Nickerson et al. (2013), each dimension in a taxonomy must contain characteristics that are mutually exclusive and collectively exhaustive. The mutually exclusive rule means that “no object can have two different characteristics in a dimension” (Nickerson et al., 2013, p. 341). We classified data marketplaces under the assumption that each data marketplace has one business model. Thus, each data marketplace is classified in one business model characteristic per dimension.

Table 7: Business model taxonomy for data marketplaces in the B2B automotive industry

	Component	Dimension	Characteristics				
Value creation	Customer segment	Domain	Location (TT, IN, HE)	Automotive (CR)	All industries (IOTA, OP)		
		Participants	Data sellers, data buyers, internal & external developers (TT,IN, HE)	Data sellers, data buyers & external developers (CR, IOTA, OP)			
	Value proposition	Data service	Customized map service (TT, IN)	Data brokering service (CR, IOTA, OP)	Both (HE)		
		Data output	Aggregated data (TT, IN)	Standardized data (CR, IOTA, OP)	Both (HE)		
		Data quality	Reviews by marketplace (TT, IN)	User reviews (IOTA, OP)	No info (HE, CR)		
		Privacy	Anonymized (TT, IN, CR)	Encrypted (HE, IOTA, OP)			
	Customer relationship	Contract	Negotiated (TT, IN, CR)	Standardized (IOTA, OP)	Both (HE)		
Value delivery	Key channels	Platform access	Closed (TT, IN, CR)	Open (HE, IOTA, OP)			
	Key resources	Platform infrastructure	Centralized (TT, IN, HE, CR)	Decentralized (IOTA, OP)			
	Key activities	Data processing activities	All (TT, IN, HE)	Limited (CR, IOTA, OP)			
Value capture	Revenue streams	Revenue streams	Usage based (TT, IN)	Usage based & freemium (HE)	Commission (CR)	Donations (IOTA)	No info (OP)
	Pricing model	Data pricing mechanism	Set by data marketplace owner (TT, IN)	Set by data seller (CR, IOTA, OP)	Both (HE)		
		Payment currency	Fiat currency (TT, IN, HE, CR)	Cryptocurrency (IOTA, OP)			

Value Creation

We specified the customer segments of data marketplaces with the dimensions *domain* and *participants*. The *domain* shows in what market the data marketplace is active. Its characteristics are location, automotive and all industries. TomTom, INRIX and HERE are specialists in the *location* domain to design dynamic maps and communicate real-time road conditions to their customers. They envision the realization of an autonomous and connected world through location technology. Caruso focuses on the complete *automotive* domain. This domain includes numerous segments such as the (i) vehicle position, movement and surroundings (ii) vehicle health and maintenance (iii) vehicle non-powertrain hardware (iv) vehicle powertrain resources (v) vehicle powertrain hardware (vi) mobility services and (vii) auxiliary devices (CR-3). With a focus on the automotive domain, Caruso aims to support the digital transformation in the entire automotive industry. IOTA and Ocean Protocol focus on *all industries*. They do not limit themselves to only the automotive industry. These data marketplace owners aim to accelerate Internet of Things (IoT) and Artificial Intelligence (AI) development by facilitating data trade across industries.

The dimension *participants* refers to the actors who are matched at a data marketplace to trade data. The data sellers and buyers who are targeted are fairly similar among the data marketplaces. All data marketplace owners target companies such as Original Equipment Manufacturers (OEMs), tier 1 suppliers and start-ups to be both data seller and data buyer at their platform. The data marketplaces differ in terms of *internal* and *external developers*. TomTom, INRIX and HERE have internal developers who use the data traded at their data marketplace to develop their value proposition. Caruso, IOTA and Ocean Protocol do not process the data internally, but target external developers who further process the datasets themselves. In these data marketplaces, the roles of marketplace owner and third party service provider are separated.

The value proposition consists of the dimensions *data service*, *data output*, *data quality* and *privacy*. In the *data service* we specify what service the data marketplace owner offers to their participants. TomTom and INRIX provide a *customized map service*. They aggregate the data from their data sellers. They map the data to navigate cars. Their service comprises real-time traffic updates, directions to charging stations for electrical vehicles, information about available parking spots and speed camera alerts. Caruso, IOTA and Ocean Protocol perform a *data brokering service*. This service comprehends minimal interference of the data marketplace owner. They do not change the content of the data from their participants, but standardize the data into one format. The data marketplace provides the technical infrastructure for direct trade between the data seller and buyer. In addition, the data marketplaces provide contracts to ensure secure data trade. HERE offers *both* the customized map service and data brokering service. These services comprehend two different value propositions.

The *data output* shows what data the data marketplace owner trades. TomTom and INRIX trade *aggregated data*, which they produce with their customized map service. Caruso, IOTA and Ocean Protocol trade *standardized data*. The data can be standardized by the data marketplace owner, as Caruso does, or the data sellers have to standardize the themselves, as IOTA and Ocean Protocol implement. HERE offers *both* aggregated and standardized data output. The output of their customized map service is aggregated data. The standardized data is the output of their data brokering service

Data quality entails who controls and preserves the data quality from the data seller. The identified characteristics are *reviews by the marketplace*, *user reviews* and *no information*. TomTom and INRIX ensure high quality data by reviewing the data themselves. Other data marketplaces such as IOTA and Ocean Protocol are not directly involved in preserving the data quality, but let their participants review the data quality. To incentivize participants to review data, the participants receive a reward. HERE and Caruso claim to provide high quality data, but we could not find information about who reviews the data quality.

Privacy indicates how stored data at a data marketplace is protected. All data marketplace owners guard data privacy by *anonymizing* (TomTom, INRIX, Caruso) and *encrypting* the data (HERE, IOTA, Ocean Protocol).

The *contracts* that data marketplace owner manages define the agreement that enforces data trade between the data seller and data buyer. INRIX, TomTom and Caruso have *negotiated* contracts with their participants to trade data. They negotiate contracts bilaterally, which results in separate agreements with their participants. The data sellers determine what data is delivered and for what purpose the data may be processed. Overall, negotiated contracts demand close partner communication, which requires high effort from the data marketplace owner. The data marketplace owner communicates with each data seller and user to understand their needs and offer personal assistance. IOTA and Ocean Protocol apply *standardized* contracts. They make use of smart contracts that operate on distributed ledger technology (DLT). Transactions are automatically updated in smart contracts, which minimizes the interaction with an intermediary party. This should decrease transaction costs at the data marketplace, which makes it suitable for multilateral trade. HERE offers *both* the negotiated and standardized contracts to trade data at their platform.

Value Delivery

The *platform access* of a data marketplace defines the degree of openness for participants to enter the platform. The open or closed platform is the channel used by data marketplace owners to reach customers with their value proposition. TomTom, INRIX and Caruso have *closed* platform access. They restrict access to their platform with identity and access management. Their participants must authenticate themselves with company details and specifications about their data use. The platform owner approves or declines the registration requests. HERE, IOTA and Ocean Protocol have *open* platform access and allow anyone to upload and buy data from the marketplace. Users can directly enter the data marketplace after they created a user account. Registration does not need to be approved by the data marketplace owner.

The *platform infrastructure* specifies how data is stored at the data marketplace. TomTom, INRIX, HERE and Caruso have a *centralized* platform infrastructure and store data in the cloud, a central location. The cloud is easily linked to clouds of other companies to exchange data. IOTA and Ocean Protocol have a *decentralized* platform infrastructure and store data across locations. There is no central administrator who controls the data. Ocean Protocol and IOTA aim to increase data transparency and data sovereignty by deploying a decentralized platform infrastructure.

The data marketplace owners perform *data processing activities* to add value to data. The main data processing activities are data collection, standardization, cleansing, storage, analysis and distribution. TomTom, INRIX and HERE perform *all* of these activities for their participants. Caruso, IOTA and Ocean Protocol perform a *limited* number of activities. They do not clean the data and are limited in data analysis. They only analyze their platform usage patterns and do not analyze the data content that is traded between their participants.

Value Capture

The *revenue streams* indicate how the data marketplace owner generates turnover. TomTom and INRIX receive *usage based* revenue streams. They create turnover by charging their participants for the usage of their data. HERE combines the *usage based and freemium* model. Participants can get up to 250.000 data transactions per month for free. When this limit is exceeded, HERE charges $45 per month for an add-on subscription or $449 per month for a pro-subscription. Caruso leaves a *commission* of the data that their data sellers sell at the marketplace. IOTA is a non-profit organization who provides their platform for free. The organization is funded by donations from individuals and enterprises to maintain their platform. Ocean Protocol does not provide information about their revenue streams.

The *data pricing mechanisms* indicate how the trading entities establish prices of the data they trade. TomTom and INRIX sell their own data for which *they set the price themselves*. At Caruso, IOTA and Ocean Protocol *the data sellers set the price* for their data that is traded. HERE is a data marketplace that applies *both* pricing mechanisms. They set the price for their own aggregated data and their data sellers set the price for the standardized data that is traded. These pricing mechanisms are examples of fixed pricing mechanisms. We did not observe dynamic pricing mechanisms at the data marketplaces we researched.

The *Payment currency* is the currency in which the payment is transferred. TomTom, INRIX, HERE and Caruso use *fiat* currency. When fiat currencies are maintained, data can be traded in multiple currencies. In the Netherlands it is the Euro, in the United Kingdom the Pound is used and in the United States the data marketplaces trade using the Dollar as currency. Ocean Protocol and IOTA have their own *cryptocurrency*. These cryptocurrencies are called tokens, and can be used only at their marketplace.

5.3 Business Model Archetypes

We grouped data marketplaces with similar business model characteristics in business model archetypes. Archetypes are reoccurring patterns in the combinations of taxonomy characteristics (Oberländer et al., 2019). We found patterns in the dimensions domain, data service and output, data quality, privacy, contract, platform access, platform infrastructure and data pricing mechanism. The characteristics of these dimensions show how data marketplaces capture, deliver and create value in distinctive manners.

The *aggregating data marketplace* owner performs data analyses as part of their customized map service to provide aggregated data. This is similar to the data marketplace owner of *the aggregating data marketplace with an additional brokering service* who also offers a customized map service. The customized map service is their core business and in addition they perform a data brokering service to enable standardized data trade between their participants. *The consulting data marketplace* owner performs a data brokering service which distincts itself from the other archetypes, because the service is paired with personal assistance of the data marketplace owner through bilaterally negotiated contracts. *The facilitating data marketplace* owner focuses on their data brokering service that runs on a decentralized platform infrastructure. This results in four business model archetypes presented in table 8. In line with our assumption that each data marketplace has one business model and the rule of mutual exclusivity, each data marketplace matches one of the business model archetypes. We further discuss these archetypes in the subsequent sections.

Table 8: Business model archetypes

Archetype	**Aggregating data marketplace**	**Aggregating data marketplace with additional brokering service**	**Consulting data marketplace**	**Facilitating data marketplace**
Case	*TomTom and INRIX*	*HERE*	*Caruso*	*IOTA and Ocean Protocol*
Orientation	*Hierarchical*	*Mixed hierarchical/market*	*Mixed hierarchical/market*	*Market*
Ownership	*Private*	*Consortium*	*Consortium*	*Independent*
Domain	Location	Location	Automotive	Cross-industry
Data service and data output	Customized map service Aggregated data	Both customized map service and data brokering service Both aggregated data and standardized data	Data brokering service Standardized data	Data brokering service Standardized data
Data quality	Reviews by data marketplace owner	Reviews by data marketplace owner	No info	Reviews by users
Privacy	Anonymized	Encrypted	Anonymized	Encrypted
Contract	Negotiated contract	Both negotiated and standardized contract	Negotiated contract	Standardized contract
Platform access	Closed	Open	Closed	Open
Platform infrastructure	Centralized	Centralized	Centralized	Decentralized
Data pricing mechanism	Set by data marketplace owner	Both set by data marketplace owner or data seller	Set by data seller	Set by data seller

Aggregating Data Marketplace

TomTom and INRIX apply the *aggregating data marketplace* archetype. They create value for their customers by aggregating the data from their data sellers to provide tailored maps for their customers. Data marketplace owners establish *personal customer relationships* with the data marketplace participants through *bilaterally negotiated contracts*. They have close contact with their participants during bilateral negotiations to understand and define data trading conditions. Data marketplace owners personally assist their customers during data collection. Osterwalder & Pigneur (2010) introduce personal assistance as a manner for business owners to build customer relationships through human interaction. Although the personal interaction demands investment from the data marketplace owner, the creation of personal customer relationships should increase the commitment of customers to their data marketplace. Moreover, the data marketplaces have *well-understood customer segments* in the *location* domain. Within the location domain, the data marketplace owner knows who their participants are, where the data comes from, what information it contains and what purpose the data is used for. Their customer groups are segmented. For example, the automotive segment comprises OEMs and Tier 1 suppliers and the enterprise segment comprehends mobile application developers, cloud providers and fleet managers. Segmented customers have slightly different needs and problems and receive differing value propositions (Osterwalder & Pigneur, 2010). This leads to the *customized value proposition* that the data marketplace owner creates by offering a *customized map service*. The data marketplace owner combines a real-time traffic service, EV service, parking service and speed camera service to create customized maps. As such, an OEM receives aggregated real-time traffic information (RTTI) and parking information for their navigation system while an external developer receives EV data to develop their own charging application. The data quality is assured by the data marketplace owner who reviews and cleans data. The data marketplace owner handles the payments, contracts and provides the infrastructure for all participants to satisfy their needs.

The aggregating data marketplace has *closed platform access*. The data marketplace owner needs to approve data seller or buyer registration before data can be sold or bought from and to the data marketplace. This contributes to a controlled environment to which participants can be denied. Furthermore, the aggregating data marketplaces need a *centralized platform infrastructure*. The centralized platform infrastructure is connected to the customer IT systems and realizes a central access point for the data marketplace owner to modify the data and perform their service.

At data marketplaces of the aggregating data marketplace archetype, *the data marketplace owner sets the price of the traded data*. The aggregated data output is owned and sold by the data marketplace owner. The usage-based data that is sold leads to direct revenue streams for the data marketplace owner.

Aggregating Data Marketplace with Additional Brokering Service

HERE applies the *aggregating data marketplace with an additional brokering service* archetype. This archetype includes two distinct value propositions. One value proposition is similar to the value proposition of the aggregating data marketplace. Data marketplace owners of both archetypes focus on delivering a customized value proposition and aggregated data within the location domain. However, data marketplace owners who apply the *aggregating data marketplace with an additional brokering service* archetype offer a second, standardized value proposition which is the data brokering service. This service enables standardized data trade directly between data sellers and data buyers at the data marketplace. The data marketplace owner uses negotiated contracts for their customized value proposition and standardized contracts for their standardized value proposition. The standardized contract enables automated assistance. Automated assistance has lower costs than personal assistance and can handle a large number of users (Osterwalder & Pigneur, 2010). The application of both negotiated and standardized contracts enables the data marketplace owner to offer personal assistance to some customers while simultaneously serving many other participants through automated assistance.

The aggregating data marketplace with an additional brokering service has *open platform access*. Anyone who creates a user account can enter the platform. The data marketplace owner deploys a *centralized platform infrastructure*. Central data storage is required for the data marketplace owner to perform data collection, standardization, cleansing, storage, analysis and distribution and deliver the customized value proposition.

To capture value, the data marketplace owner maintains two data pricing mechanisms. The *data marketplace owner sets the price* for the aggregated data that is produced with the customized map service and the *data sellers set the price* for the standardized data that they sell via the brokering service.

Consulting Data Marketplace

Caruso applies the *consulting data marketplace* archetype. They offer a *standardized value proposition*, as does the data marketplace owners who apply the *aggregating data marketplace with additional brokering service* archetype. Significant for the brokering service of the data marketplace with *consulting data marketplace* archetype is that the data marketplace owner pairs the service with negotiated contracts. The data marketplace owner negotiates the contract conditions with their participants bilaterally. The data marketplace owner gains knowledge about the data needs and price preference of their participants and aligns the needs of their data sellers and data buyers. If a data seller wants to sell specific data assets at the marketplace, there needs to be a data buyer interested in buying those segments and vice versa. The participants are personally assisted on bilateral basis by the data marketplace owner through negotiated contracts. Similar to the contracts of the aggregating data marketplace, these contracts lead to *strong customer relationships*. The customers are identified in the automotive domain. The data marketplace owner aims to serve all participants with an interest in automotive data. Potential participants are OEMs, any supplier of the OEM, insurance companies, infotainment services and external developers. The consulting data marketplace is the intermediary who connects these interdependent groups.

The consulting data marketplace has *closed platform access*. Participants can only enter the platform after they are provided login credentials by the data marketplace owner. This provides controlled provision and purchase of data at the marketplace. Furthermore, consulting data marketplaces have a *centralized platform infrastructure*. The data marketplace owner stores and publishes metadata about the datasets in the centralized platform infrastructure. The metadata is analyzed to create insights about the platform usage patterns. Significant for the data marketplace with the consulting data marketplace archetype is that they do not store the exchanged data sets in their cloud, but only keep track of metadata about the datasets.

The consulting data marketplace allows *the data seller to determine the price* of the sold data. The data marketplace owner consults their participants about possible data pricing mechanisms. The revenue streams for the exchanged data are transferred between the data seller and buyer. The data marketplace owner receives a commission of the sold data from the data seller and is paid for their provided service.

Facilitating Data Marketplace

IOTA and Ocean Protocol apply the *facilitating data marketplace* archetype. They coordinate transactions between data sellers and buyers through the data brokering service without interference of the data marketplace owner. The facilitating data marketplace contains a *standardized value proposition* that comprises a data brokering service. The data marketplace owner aims to provide access to data that participants did not have access to before to further develop IoT and AI technologies. Developers lack data to improve their algorithms and larger companies lack advanced algorithms to analyze their data. At the data marketplace, these participants can trade data across *all industries*. The data marketplace participants process the standardized data and review the data quality themselves, with minimal interference of the data marketplace owner. The data marketplace owner does not offer personal assistance like the data marketplace owner who applies the *consulting data marketplace* archetype, but uses standardized, smart contracts. This foresees a high number of transactions between participants and *automizes the process* of data trade.

The facilitating data marketplace has *open platform access*. Anyone who knows how to use the infrastructure and has a need to trade data can join the ecosystem. In addition, the facilitating data marketplace is the only business model archetype that includes a *decentralized platform infrastructure*. The DLT is the building block that facilitates the value proposition of the data marketplace owner. The decentralized platform infrastructure allows for minimal intervention of the data marketplace owner and direct transactions between the data seller and buyer. Transactions in DLTs are immutable and transparent, to ensure safe data delivery. The main task of the data marketplace owner is to define transaction rules and link transactions to be executed and verified by the participants.

The marketplace owners who apply the facilitating data marketplace archetype enable *the data sellers to set the price* for the traded datasets. The revenue streams are directly transferred between the data seller and data buyer. These data marketplaces are owned by non-profit organizations. They do not intend to make profit from the data that is traded at their platform.

6. DISCUSSION

Our business model archetypes are distinctive for the data marketplace types defined in section 4. TomTom and INRIX, the data marketplace types with private ownership and a hierarchical orientation, apply the *aggregating data marketplace* archetype. HERE and Caruso, data marketplace types with consortium ownership and characteristics from both a hierarchical and market orientation, apply the *aggregating data marketplace with additional brokering service* and *consulting data marketplace* archetypes. IOTA and Ocean Protocol, data marketplace types with independent ownership and a market orientation, apply the *facilitating data marketplace* archetype.

A value proposition that offers a solution instead of data 'items' In its essence, all data marketplaces trade data. Through performing additional services in their value proposition, a data marketplace owner can distinguish their marketplace from other data marketplaces. Our business model archetypes show that data marketplace owners create additional value for their customers by performing a customized map service, reviewing the data quality or offering personal assistance through negotiated contracts. The value proposition of data marketplaces with the *facilitating data marketplace* archetype is the only value proposition that focuses solely on a data brokering service.

The value proposition of the *facilitating data marketplace* represents the problem that Teece (2010) describes as the sale of 'items' instead of the sale of a solution. Data assets, or 'items', could be described as 'intangibles', 'know-how' and 'technological components'. These goods are difficult to price and are rarely traded in market structures (Koutroumpis et al., 2017; Powell, 1990; Teece, 2010). According to Teece (2010), it is a common problem that the sale of assets that do not have perfect property rights, leads to market failure. Business owners who apply business models that are based on selling intangibles may not capture significant value with their value proposition. Therefore, companies who trade intangible assets need to bundle them into a solution.

The *aggregating data marketplace, aggregating data marketplace with additional brokering service* and *consulting data marketplace* archetypes comprise value propositions in which data is bundled into a solution. The data marketplace owners of these archetypes trade data and provide complementary services such as a customized map service, data quality reviews or personal consultation about data sale and purchase. Spiekermann (2019) argues that the performance of such services is a key success factor in the business model of data marketplaces because it increases value for the customer. He finds that data marketplace owners who aggregate data or assure data quality, create value as they go beyond data forwarding (Spiekermann, 2019). The performance of such services does require higher investment in time and money from the data marketplace owner. Osterwalder & Pigneur (2010) explain that these companies focus on delivering a premium value proposition and have value-driven business models. Their customers do not only pay for the data that they get, but also for the service that the data marketplace owner performs. The data marketplace owners can sell their solution against a higher price, which their customers are willing to pay for.

Data marketplace owners who apply the *facilitating data marketplace* archetype focus on data forwarding with their brokering service. These data marketplace owners have a lean cost structure and automize most of their processes. This is what Osterwalder & Pigneur (2010) call a cost-driven business model. Data marketplace owners who apply this business model promise an increase in data accessibility for their participants against a low price. Their value proposition entails trade in data 'items'. However, this does not appear to be the solution for their customers. Data sellers and buyers remain absent, which diminishes their ability to increase data access. There is a need for data marketplace owners from the *facilitating data marketplace* archetype to bundle their data brokering service with complementary services. This way they can attract data sellers or buyers by offering a solution instead of trading data 'items'.

The establishment of strong customer relationships or competitive pricing Data marketplace owners build customer relationships to attract customers and sell their value proposition. In our archetypes, we recognize that data marketplace owners who apply the *aggregating data marketplace*, *aggregating data marketplace with additional brokering service* and *consulting data marketplace* archetypes implement bilaterally negotiated contracts. As explained in section 5.3, data marketplace owners who personally assist their customers in bilateral negotiations, build personal customer relationships. This aligns with results of Koutroumpis et al. (2020) who note that one-to-one data marketplaces have "relational contracts" which are long term and enable repeated interaction between the data marketplace owner and their participants. We expect repeated interaction in organizations with a hierarchical trading structure. As Powell (1990) explains, the personal identification between the trading parties in a hierarchy causes them to trade repeatedly with each other. Actors who trade in these organizations are driven by routines and have less room to display opportunistic behavior (Powell, 1990; Williamson, 1973).

On the contrary, actors in organizations with a market structure aim to minimize their personal costs and behave opportunistically (Williamson, 1973). The buyers easily switch between sellers when they are not satisfied by certain pricing conditions. As Powell (1990) explains, price competition highly influences the behavior of actors in hierarchies. The trading parties seek quick and efficient interactions. Koutroumpis et al. (2017) who find that many-to-many data marketplace owners standardize contract conditions to increase efficiency and lower transaction costs. We find similar results. Owners of data marketplaces with a market orientation apply the *facilitating data marketplace* archetype. This archetype includes standardized contracts. The data marketplace owners implement standardized contracts to offer their customers automated assistance which is efficient and has lower costs than personal assistance.

To satisfy the needs of the actors that data marketplaces with a market orientation attract, they need to set a competitive environment and keep product prices low. This requires data marketplace owners who apply the *facilitating data marketplace* archetype to have dynamic pricing mechanisms and high numbers in demand and supply. However, the high number of data sellers and data buyers has not yet been reached at data marketplaces with a market orientation. These data marketplace types fail or remain in the conceptual phase (Koutroumpis et al., 2017; Spiekermann, 2019). As shown in our business model archetypes, dynamic pricing mechanisms do not occur either. Instead, fixed data pricing mechanisms, set by the data marketplace owner or data seller, are applied in practice. Fruhwirth et al. (2020), who researched 20 data marketplaces, found that 2 data marketplace owners establish prices based on auction or negotiation. The other data marketplaces they researched have fixed pricing mechanisms. Out of the 16 data marketplaces that Spiekermann (2019) researched, only 4 data marketplace owners priced data based on market supply and demand. One of those data marketplaces withdrew from the market and the others are still in the conceptual stage. The expected functioning of the invisible hand of the market remains obsolete. Because a competitive environment is not established, data marketplaces with a market orientation fail to attract participants who trade on competitive basis and aim for the maximum individual gain.

7. CONCLUSION AND FUTURE RESEARCH

We analyzed the business models of different types of data marketplaces that range from hierarchical to market orientation and private to independent ownership in the B2B automotive industry. The research question that we address states: *What business model archetypes are applied by data marketplace owners from different types of data marketplaces in the B2B automotive industry?*

In this paper, a data marketplace is defined as an organization with a hierarchical or market orientation and private, consortium or independent ownership that matches buyers and sellers, facilitates transactions and provides an institutional infrastructure to trade machine-readable data. We created a taxonomy in which the business models of TomTom, INRIX, HERE, Caruso, IOTA and Ocean Protocol are specified. The characteristics of thirteen dimensions distinguish one business model from the other. Patterns are recognized in the dimensions domain, data service, data output, contract, platform access, platform infrastructure and data pricing mechanism. This results in four business model archetypes which are linked to the data marketplace types.

TomTom and INRIX are data marketplaces with private ownership and a hierarchical orientation. They apply the **aggregating data marketplace** archetype and process data from their sellers to aggregate data into a customized value proposition. HERE is a data marketplace with consortium ownership and characteristics from both the hierarchical and market orientation. They apply the **aggregating data marketplace with additional brokering service** archetype. HERE aggregates data to create a customized value proposition as core business and provides an additional data brokering service as additional standardized value proposition. Caruso belongs to the same data marketplace type as HERE and applies the **consulting data marketplace** archetype. They provide a data brokering service and advise their participants about the usage and exchange of their data. IOTA and Ocean Protocol are data marketplaces with independent ownership and a market orientation. They apply the **facilitating data marketplace** archetype and deploy a decentralized platform infrastructure to coordinate transactions between data sellers and buyers with their data brokering service.

The owners of data marketplaces with a market orientation and independent ownership, which are conceptual, apply the facilitating data marketplace archetype. This archetype is not proven effective in practice yet. The other business model archetypes are applied by owners of data marketplaces that are past the conceptual stage. Those business model archetypes are effective for data marketplaces in practice.

In the emerging research field of data marketplaces, few taxonomies are developed to structure business models of data marketplaces. The taxonomies of Spiekermann (2019) and Fruhwirth et al. (2020) cover one type of data marketplace; data marketplaces with a market orientation and independent ownership. These data marketplaces are conceptual. We contribute to academic knowledge by including data marketplace types ranging from hierarchical to market orientation and private to independent ownership. The inclusion of these data marketplace types enables the identification of business models that data marketplace owners actually apply in practice. Our definition of data marketplaces in a market–hierarchy continuum may be controversial as some researchers could argue that all data marketplaces must have a market orientation. However, we deemed it necessary to include the hierarchical orientation to research business models of data marketplaces in practice. We advise future researchers to continue making a distinction between data marketplaces with a hierarchical orientation and market orientation when researching data marketplaces in practice.

Our taxonomy offers a starting point for other researchers to further structure business models of data marketplaces and identify new business model dimensions and characteristics. Our taxonomy, based on the B2B automotive industry, can be extended in two ways. On the one hand, additional data marketplaces from the B2B automotive industry can be classified. During the evaluation interviews Otonomo and oneTRANSPORT were suggested as additional data marketplaces in the B2B automotive industry. The classification of these additional data marketplaces may result in a more reliable and exhaustive taxonomy for the automotive industry. On the other hand, more data marketplaces from industries different than the automotive industry can be classified. The insurance industry is for example mentioned by Koutroumpis et al. (2017) as an industry with data marketplaces past the conceptual phase. However, the classification of data marketplaces from other industries may require a more generic interpretation of business model characteristics to enable comparison among cases. The researcher has to make a trade-off between the conciseness of the taxonomy and the granularity in business model characteristics.

We did not evaluate our taxonomy with data marketplace practitioners due to time constraints. In cooperation with the experts who designed the existing taxonomies we speculated that our taxonomy could be useful for data marketplace owners. Data marketplace owners can make design choices based on our taxonomy to develop their data marketplaces. For instance, practitioners who are still designing their data marketplace can use our taxonomy to select characteristics for their own business model. Furthermore, practitioners from data marketplaces that are past the conceptual stage can use our taxonomy for a competitor analysis. They can identify whether their competitors are innovating their business models in areas where they should evolve as well. Whether our taxonomy suffices for those purposes is not evaluated and is advised for future research.

REFERENCES

Agahari, W., de Reuver, M., & Fiebig, T. (2019). Understanding how privacy-preserving technologies transform data marketplace platforms and ecosystems: The case of multi-party computation.

Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. European Journal of Information Systems, 19(3), 359–376. https://doi.org/10.1057/ejis.2010.21

Amit, R., & Zott, C. (2001). Value creation in e-business. Strategic Management Journal, 22(6–7), 493–520. https://doi.org/10.1002/smj.187

Bakos, Y. (1998). The Emerging Role of Electronic Marketplaces on the Internet. Communications of the ACM, 41(8), 35–42. https://doi.org/10.1145/280324.280330

Charmaz, K. (2006). Constructing grounded theory: a practical guide through qualitative analysis. Sage.

Cheng, B., Longo, S., Cirillo, F., Bauer, M., & Kovacs, E. (2015). Building a Big Data Platform for Smart Cities: Experience and Lessons from Santander. Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress 2015, 592–599. https://doi.org/10.1109/BigDataCongress.2015.91

Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. Industrial and Corporate Change, 11(3), 529–555. https://doi.org/10.1093/icc/11.3.529

Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Platforms and infrastructures in the digital age. Information Systems Research, 29(2), 381–400. https://doi.org/10.1287/isre.2018.0794

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering Business Models of Data Marketplaces. 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? Cambridge Service Alliance Blog. http://cambridgeservicealliance.blogspot.co.uk/2014/04/big-data-for-big-business_3.html

Keller, R., & König, C. (2014). A reference model to support risk identification in cloud networks. 35th International Conference on Information Systems, ICIS 2014.

Koutroumpis, P., Leiponen, A., & Thomas, L. (2017). The (Unfulfilled) Potential of Data Marketplaces. ETLA Working Papers, 2420(53). http://pub.etla.fi/ETLA-Working-Papers-53.pd

Koutroumpis, P., Leiponen, A., & Thomas, L. D. W. (2020). Markets for data. Industrial and Corporate Change, March. https://doi.org/10.1093/icc/dtaa002

Mao, W., Zheng, Z., & Wu, F. (2019). Pricing for Revenue Maximization in IoT Data Markets: An Information Design Perspective. Proceedings - IEEE INFOCOM, 2019-April, 1837–1845. https://doi.org/10.1109/INFOCOM.2019.8737571

Martens, B., & Mueller-langer, F. (2018). Access to digital car data and competition in aftersales services markets. EC JRC Digital Economy Working Paper 2018-06.

Mišura, K., & Žagar, M. (2016). Data marketplace for Internet of Things. Proceedings of 2016 International Conference on Smart Systems and Technologies, SST 2016, 255–260. https://doi.org/10.1109/SST.2016.7765669

Möller, F., Bauhaus, H., Hoffmann, C., Niess, C., Otto, B., & Isst, F. (2019). Archetypes of Digital Business Models in Logistics Start-Ups. ECIS 2019 Proceedings, 1–18.

Muschalle, A., Stahl, F., Löser, A., & Vossen, G. (2012). Pricing approaches for data markets. International Workshop on Business Intelligence for the Real-Time Enterprise, 129–144. https://doi.org/10.1007/978-3-642-39872-8_10

Musso, S., Perboli, G., Rosano, M., & Manfredi, A. (2019). A Decentralized Marketplace for M2M Economy for Smart Cities. 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), i, 27–30. https://doi.org/10.1109/wetice.2019.00014

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. https://doi.org/10.1057/ejis.2012.26

Oberländer, A. M., Lösser, B., & Rau, D. (2019). Taxonomy Research in Information Systems: a Systematic Assessment. Proceedings of the 27th European Conference on Information Systems (ECIS), 0–17. https://aisel.aisnet.org/ecis2019_rp/144

Osterwalder, A., & Pigneur, Y. (2010). Business model generation : A handbook for visionaries, game changers, and challengers.

Park, J. S., Youn, T. Y., Kim, H. Bin, Rhee, K. H., & Shin, S. U. (2018). Smart contract-based review system for an IoT data marketplace. Sensors (Switzerland), 18(10), 1–16. https://doi.org/10.3390/s18103577

Perera, C., Wakenshaw, S. Y. L., Baarslag, T., Haddadi, H., Bandara, A. K., Mortier, R., Crabtree, A., Ng, I. C. L., McAuley, D., & Crowcroft, J. (2017). Valorising the IoT Databox: creating value for everyone. Transactions on Emerging Telecommunications Technologies, 28(1). https://doi.org/10.1002/ett.3125

Powell, W. W. (1990). Neither Market Nor Hiearchy: Network forms of organization. Reseach in Organizational Behavior, 12(January), 295–336.

Schomm, F., Stahl, F., & Vossen, G. (2013). Marketplaces for Data : An Initial Survey. ACM SIGMOD, 42(1), 15–26.

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. Intereconomics, 54(4), 208–216. https://doi.org/10.1007/s10272-019-0826-z

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. Vietnam Journal of Computer Science, 3(3), 137–143. https://doi.org/10.1007/s40595-016-0064-2

Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Because your taxonomy is worth it : Towards a framework for taxonomy evaluation. Proceedings of the Twenty-Seventh European Conference on Information Systems (ECIS), Stockholm, April, 0–19.

Teece, D. J. (2010). Business models, business strategy and innovation. Long Range Planning, 43(2–3), 172–194. https://doi.org/10.1016/j.lrp.2009.07.003

Thomas, L. D. W., & Leiponen, A. (2016). Big data commercialization. IEEE Engineering Management Review, 44(2), 74–90. https://doi.org/10.1109/EMR.2016.2568798

Tiwana, A. (2013). Platform Ecosystems: Aligning Architecture, Governance and Strategy. Morgan Kaufmann.

Williamson, O. E. (1973). Markets and Hierarchies: some elementary considerations. The American Economic Review, 63(2), 316–325.

Williamson, O. E. (1989). Chapter 3 Transaction cost economics. Handbook of Industrial Organization, 1, 135–182. https://doi.org/10.1016/S1573-448X(89)01006-X

B. Data Ecosystem

Data marketplaces trade cross-industry or specialize in one specific industry (see Figure 20). As argued in section 1.3, this research focuses on the automotive industry. This appendix provides an understanding of stakeholders and data sources involved in automotive data marketplaces.

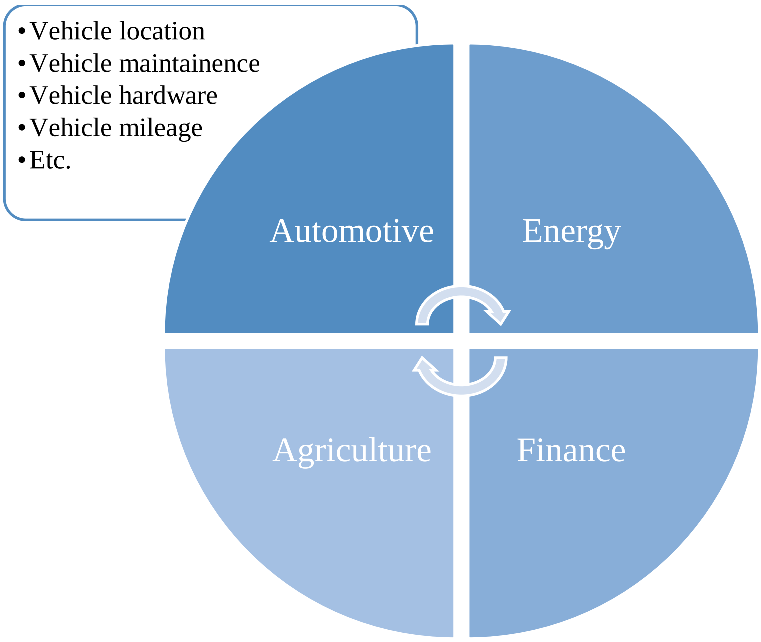

Figure 20: Data ecosystem

B.1 Stakeholders

In relation to business models, stakeholders are important to gain insights in the customer segment and key partners for data marketplaces. Overall, recognized stakeholders at data marketplaces are data seller, data buyer, third party service provider and data marketplace owner (Koutroumpis et al., 2017; Spiekermann, 2019). Muschalle et al. (2012) identify data marketplace stakeholders and their roles on a more granular level. First, the data market owner develops the data marketplace to store, search and exchange data. Second, data providers utilize data marketplaces to store and advertise data. Third, analysts are domain experts who use data exploration to compose meaningful reports. Fourth, application vendors process data into applications to ease data access for a broader audience. Fifth, developers of data associated algorithms develop and upload algorithms to be used by analysts and application vendors for data mining, matching, cleansing, etc. Sixth, consultants support analysts by advising on for example data source collection, data integration and product development. Seventh, licensing and certification entities offer "branded" data, applications and algorithms.

In practice, one stakeholder can fulfill multiple roles. Thomas & Leiponen (2016) note that organizations start with commercializing their own data and move to other business models in which they work with partners and suppliers to perform more activities such as data aggregation in order to sell this information to other parties. Thereby, an industrial player can move from data seller to third party service provider, combined with the role of a data marketplace owner. In such a case, the data marketplace can take on the role of the analyst, application vendor, algorithm developer, consultant and licensing entity.

There are numerous stakeholders in the automotive industry who have an interest in automotive data marketplaces. They could take multiple stakeholder positions. The following sections discuss some of stakeholders in the automotive industry and what roles they can fulfill at a data marketplace.

Car Owners

Even though car owners are classified as end consumers and the focus of this research is on B2B, the interest of car owners in data marketplaces are briefly discussed as the trade of data concerns their personal data. While car owners have full power over their choice of product or service, the data they produce by using their purchase is in control of the providing company. As such, car manufacturers have the power to sell their consumer data. The data produced by the connected cars such as mileage, speed and location is of relevance to data marketplaces trading in automotive data. When car owners want to buy products or make use of services, they often have no other choice than to accept the terms and conditions, which allows the provider to further process their data. This makes the *power* of car owners in data trade *low*. In the B2B sphere, individual car owners have *low interest* in the trade of data. Although the data contains information about their behavior, there are few use cases in which individual consumers can commercialize their own data or use the data for private purposes.

Governments

Governments have an interest to make open data widely available to stimulate innovation of both businesses and governments (Janssen & Zuiderwijk, 2014). Governmental bodies can be identified on European level, national level and regional level. On European level, the **European Commission** (EC) envisions a single market for personal data, non-personal data, public and private data (European Commission, 2020). The EC creates regulatory frameworks, such as the GDPR, which need to be followed by nations and companies. Also, the EC requires over funding to stimulate innovation. As stated in their data strategy, the EC plans to invest €2 billion to stimulate European data projects (European Commission, 2020). The regulatory and finance resources give the EC *high power* in the trade of data. **National governments** are expected to provide high quality datasets at data marketplaces (European Commission, 2020). The Dutch Ministry of Internal Affairs for example initiated a developers portal which interfaces all APIs of the Dutch government (developer.overheid.nl, 2019 May 22). In addition, national governments need to enforce European regulation. As such, national governments require over *high institutional power* as well. **Regional governments** have *less power* than the EC and national governments, as they merely execute data regulation. However, they do require over a lot of data. For example, municipalities such as The Hague manage open data portals which contain information about parking spots, air quality and speed limits. According to Bertoncello et al. (2016) governments have two main interests in automotive data marketplaces. First, they set up rules and standards for data trade as regulators. Second, they optimize the national infrastructure for the public good. Governments have a lot of information on demographics of citizens in particular cities or regions as well as geographical data (Ramirez et al., 2014). This data is often available open source and can be uploaded on data marketplaces. With data initiatives, governmental bodies can improve societal living conditions and enhance economy. Therefore, the *interest* of all governmental bodies in data trade is *high*.

OEMs

Car manufacturers, **original equipment manufacturers** (OEMs), design the car and sell it to the customer. They play a key role as a data seller since they have car data stored in their branded data silos. The lock-in effects which car manufacturers create for specific hardware and software built into their cars give OEMs power over their user data (Martens & Mueller-langer, 2018). This puts OEMs in *powerful* positions regarding data trade, because they require over valuable datasets which they have exclusive access to. Besides, OEMs are multinationals with high revenues. For example, BMW had a revenue of €104 billion in 2019 (BMW, 2019). Thus, OEMs have the budget to become data buyers and data marketplace owners if they want to. Various factors withhold large companies from sharing their data such as a lack of economic incentives, a lack of trust between businesses and imbalances in negotiating power (European Commission, 2020). While OEMs can use data marketplaces to further advance their own use cases and can make profit by selling their data, OEMs can perceive data marketplaces as a threat since it puts their exclusive power over data in danger. Therefore, the *interest* in data trade at data marketplaces is *medium*.

Automotive Suppliers

Automotive suppliers are companies who deliver software and hardware to OEMs to build the cars. These companies are often divided into tier 1, tier 2 and tier 3 suppliers. Tier 1 suppliers directly deliver parts to the OEM. Tier 2 suppliers often specialize in a specific domain and work across industries. Their products wind up in cars, but tier 2 suppliers do not have a direct connection to OEMs like tier 1 suppliers do. Tier 3 suppliers deliver raw materials which OEMs, tier 1 and tier 2 suppliers need. These suppliers use data to improve their service to end consumers and B2B customers (Bertoncello et al., 2016). An example of a **tier 1 supplier** is Bosch. Bosch intends to be a connected mobility services supplier and develops software and hardware to be used in connected vehicles. In 2019 Bosch made a revenue of €77.7 billion (Bosch, 2019). Bosch has a similar power and interest in data marketplaces as OEMs. They have the budget and data to become data sellers, buyers and data marketplace owners, but can also perceive data marketplaces as a threat to their own business.

IT infrastructure providers

IT infrastructure providers are required to build the data marketplace. A data marketplace requires cloud services for data hosting, algorithms for data processing and developer toolkits (European Commission, 2020). Companies such as Microsoft provide such resources. They are specialized in IT infrastructures and have the knowledge and funding to build data marketplaces. The Azure marketplace of Microsoft has data catalogs at which users can browse for data while Microsoft also support the IT infrastructure of other data marketplaces with their Azure cloud. Thereby, they can be a facilitator to set up data marketplaces or do it themselves. Thus, IT infrastructure providers have *high power* and *high interest* in data marketplaces.

Aftersales service providers

Aftersales service providers are companies who leverage car data to further develop their own products. Three aftersales services providers are highlighted. These are navigation services, insurance companies and other app developers. First, **navigation services** enrich their datasets with GPS information from the cars. Besides GPS data, weather and parking data contribute to a more accurate navigation service. An example of a navigation service is TomTom. They buy data from various sources and combine the data into traffic information. This data output is sold to individual developers and OEMs. Navigation services rely on the willingness of other stakeholders to share data with them. Competition between navigation services such as TomTom, HERE and INRIX is quite high. Therefore, the *power* of navigation services is *medium*. Second, **insurance companies** can offer more customized products based on a customer's driving behavior (Martens & Mueller-langer, 2018). They generally require over financial resources. For example, Allianz had a revenue of €11.9 billion (Allianz, 2020). However, the development of data marketplaces does not depend on the data provision or acquisition of insurance companies. Therefore, they have *low power* in automotive data marketplaces. Third, **app developers** can be any start-up initiative wanting to develop a new application related to car services. They can acquire new datasets which they analyze for their own purpose. Examples of use cases for app developers are the development of a parking navigator in urban cities, automated payment service for electrical vehicle charging or a delivery service to someone's car. It is crucial for aftersales service providers to gain access to automotive datasets to develop their ideas and improve their services. Their *power* is *low* in the trade of data, because they are dependent on datasets from OEMs and tier 1's which are locked-up. The *interest* is *high* because individual developers succeed when they have access to the value datasets.

B.2 Data Sources

Data marketplaces can make use of multiple key resources for the traded data. There are three main sources which can be consulted. These are governmental data sources, social media data sources and commercial data sources.

Governmental Data Sources

Governments have open data portals where demographical and geographical data can be found. Ramirez et al. (2014) list a range of information which governments can provide from professional to recreational licenses and from driving records to court records. This data is produced with state money (European Commission, 2020). Therefore, a requirement for the use of this data is that it should benefit the society. Data marketplaces acquire is data for free from most governmental data portals.

Social Media Data sources

Social media platforms such as Twitter and LinkedIn can provide user data via their APIs. If users did not restrict their privacy settings, this information is openly available. Names, locations and opinions about certain topics are examples of information which can be extracted from these sources.

Commercial Data Sources

Commercial data comes from private companies who have customer data. This information may be personal sensitive, such as health-related records, and less personal sensitive, such as vehicle purchase (Ramirez et al., 2014). This data needs to be acquired directly from the respective company for a financial compensation.

B.3 Overview

The stakeholders from appendix B.1 and the data sources from appendix B.2 are merged in Table 26. It provides an overview of potential roles which stakeholders can play in B2B automotive data marketplaces. It must be noted that the information is gathered from non-technical reports about data marketplaces and stakeholder websites. In order to fully understand the resources, power and interests of the stakeholders, further qualitative research such as conducting interviews is required.

The resources show what means stakeholders have to realize data marketplaces. The main resources are data (governmental and commercial), money (funding) and knowledge (algorithms). The power of stakeholders indicates how much influence the stakeholder has on the realization of a data marketplace. This is influenced by the resources of the stakeholders (Bouwman et al., 2008). The interest of a stakeholder concerns the extent to which the stakeholder wants to trade data at a data marketplace.

Table 26: Stakeholder analysis

Stakeholder	Example	(Potential) roles	Resources	Power	Interest
Individual car owners	Citizen	Data subject	Personal data	Low	Low
European Commission	EC	Regulator Facilitator	Regulation Funding	High	High
National government	Ministry of Internal Affairs	Data seller Regulator	Regulation Funding Governmental data	High	High
Regional government	Municipality	Data seller	Governmental data	Medium	High
OEM	BMW	Data seller Data buyer Data marketplace	Commercial data Budget	High	Medium

Tier 1 supplier	Bosch	Data seller Data buyer Data marketplace owner	Commercial data Budget	High	Medium
IT infrastructure providers	Microsoft	Data marketplace owner	Algorithms Budget	High	High
Car fleets		Data seller Data buyer	Commercial data	Low – medium	Medium
Mobile broadband platform providers	KPN	Data seller	Commercial data	Low – medium	Medium
Navigation services	TomTom	Data seller Data buyer Data marketplace owner	Algorithms Commercial data Budget	Medium	High
Insurance companies	Allianz	Data seller Data buyer	Commercial data	Low	Medium
Individual developers	MOBI	Data buyer	Algorithms	Low	High

C. Intensive Interviews

This appendix presents the analysis of the interviews in realm of the Grounded Theory, which were held to extend the existing business model taxonomies for data marketplaces. These interviews function to discover what business model components are important for data marketplaces and what dimensions should be included in the taxonomy. Over 25 employees of data marketplaces were approached via LinkedIn and Email to schedule interviews. These employees were approached based on their job title and data marketplace classification type. The aim was to talk to a diverse range of data marketplaces who can be classified as hierarchical and market data marketplaces and private and independent data marketplaces. Of the approached employees, 10 people responded with a positive answer and 7 interviews were held. More positive responses were received from the independent data marketplaces, which is why they are better represented than the private and consortium types. Table 27 shows an overview of the interviewees.

Table 27: Interview respondents

Code	Type	Job title
DM1	Market, consortium	Business development
DM2	Market, independent	Lead data scientist/product owner
DM3	Market, independent	Unknown
DM4	Market, independent	Product and Business Development
DM5	Market, government	Innovation Manager Smart Mobility
DM6	Market, private	Director Business Development
DM7	Hierarchical, private	Head of Enterprise Business Development

The interviews are conducted in the realm of the Grounded Theory in order to discover categories from a practical approach before literature is analyzed. This has the advantage to discover concepts which otherwise could have been overlooked and it contributed to the goal of adapting the taxonomy based on practical data marketplaces. Intensive interviewing is applied during the interviews to explore new perspectives. This method is recommended by Charmaz (2006) in the Grounded Theory approach to explore topics open-ended and in-depth. A number of questions are formulated to lead the interview as shown below. However, no interviews were the same as follow-up questions were asked to interviewees if topics needed more explanation and interviewees were given the freedom to give their input. The leading questions are as follows:

1. What are the main trends for data marketplaces and how does [name data marketplace] react to those trends?
2. What are the main challenges for data marketplaces and how does [name data marketplace] react to those challenges?
3. Could you explain the key components of the [name data marketplace] business model?
4. What is the difference of [name data marketplace] with other data marketplaces?

After initial and focused coding, main categories are created from the interviews. This can be perceived as a new round of focused coding during which the focused codes of various interviews are aligned into separate categories. This step is required, because interviewees use different wordings to describe the same concept. This leads to focused codes with different wordings, although they belong to the same category. For example the focused code *comply with GDPR* and *privacy regulation* as well as *smart contract* and *terms and conditions* can be aggregated into the same category *data regulation*. This process results in 7 main categories which are presented in Table 28.

Table 28: Data categories

Main categories	Focused codes
Data regulation	Design quality standards / smart contract / comply with GDPR / delegated data regulation / preserve data privacy / comply with EU law / setting legal framework is challenging / terms and conditions / privacy is a challenge / privacy regulation / privacy challenge / check privacy regulation / terms and conditions determine data usage / Privacy disables open data publication / use of data is a license / non-cooperation of OEMs / customers restrict data usage / data ownership
Customers	Users / industry domain / attracting a specific customer segment / large target group / customer segment / maintain customer segments / direct OEM relationship
Platform infrastructure	Decentral data control / open governance / centralized or decentralized approach / open protocol / open platform infrastructure / decentral infrastructure / decision making at consumer / challenge to regulate IT integration
Data processing activities	Searching databases / saving time / overview in catalog / perform additional activities / provide corporate and open data / aggregate data / perform activities for all needs / advertise meta data / mixed functionalities / preserve data privacy / national access point / offer broker services / regulate data availability / broker of data / harmonize and synchronize data / setting legal framework is challenging / data processing activities / acting as traditional marketplace / extracting value from data is a challenge / acting as a consent management hub / key activities / key processes / perform several activities / advise OEM in data supply / market leader for development and research / develop own data products / perform activities / enable analysis of car data / generating insights from sensor data is difficult / collect data which is needed / differentiate added value / expand product offering / value chain depends on layer / multiple suppliers cause more activities / production is partially standardized and partially customized / try to standardize terms and conditions / high service quality / added value of aggregated product / performance of data processing activities
Revenue model	Valorizing data is difficult / no active role in data pricing / price discovery / explaining the crypto currency is a challenge / pricing / online product prices / data licensing / Licensing disables open data publication
Data quality	No check of data quality / cooperate to create quality standards / cooperate to automatically improve data
Other	Data marketplace concept / small company size / explain data marketplace concept / role MDM / create fit between customer and marketplace / data marketplace type / fit between governance and client / increase in data generation / more hardware in vehicles / evolve mobility definition / cooperate to develop solution / More data collection because of partnerships

D. Preliminary Taxonomy

The preliminary taxonomy distinct business model components from business model dimensions. The business model components are derived from the business model canvas of (Osterwalder & Pigneur, 2010) as described in section 2.1.2. The description of these components is shown in Table 29.

Table 29: Business model components descriptions (Osterwalder & Pigneur, 2010)

Business model component	Description
Customer segment	the different groups of people or organizations served by an enterprise
Value proposition	the bundle of products and services that create value for a specific Customer Segment
Customer relationship	the types of relationships a company establishes with specific Customer Segments
Channels	how a company communicates with and reaches its Customer Segments to deliver a Value Proposition
Key resources	the most important assets required to make a business model work
Key activities	the most important things a company must do to make its business model work
Key partners	the network of suppliers and partners that make the business model work
Revenue streams	the money a company generates from each Customer Segment
Data pricing mechanism	Each Revenue Stream has different pricing mechanisms. The main pricing mechanisms are fixed and dynamic pricing
Cost model	all costs incurred to operate a business model

The preliminary taxonomy includes numerous business model dimensions which are derived from following the Grounded Theory method. After combining the interview results with technical and non-technical sources, several business model dimensions are included in the taxonomy. These dimensions are described in Table 30.

Table 30: Preliminary business model dimension descriptions

Dimension	Description
Domain	the category which the offered datasets stem from
Participants	the data sellers and buyers who are matched at the data marketplace
Privacy	the privacy protection of the stored data
Data source	the source where the data on the data marketplace is collected from
Data output	the transformed state in which the data product is delivered
Data quality	the quality guarantee of the purchased data
Contract	the agreement which enforces the data trade
Platform access	the access terms to the platform, through which the value proposition is delivered
Platform infrastructure	the storage location of data
Data processing activities	the activities performed by the data marketplace owner which increase the value of the data
Key partners	the alliances created to optimize the business model, reduce risk or acquire resources
Revenue streams	the way in which the data marketplace owner generates turnover by charging fees to its customers for a data transaction, marketplace membership, listing of data product, storage space or use of value-adding services
Data pricing mechanism	The pricing mechanism that is used to establish the price of the data output that is traded between the trading entities
Payment currency	the currency in which the payment is transferred

The selected business model components, dimensions and characteristics are shown in the preliminary taxonomy in

	Component	Dimension	Characteristic	
Value creation	Customer segment	Domain	To be defined	
		Participants	To be defined	
		Privacy	Anonymized	Encrypted

	Value proposition	Data source	Government	Social media	Commercial	Self-generated	Community	
		Data output	Transformed data	Non-transformed data				
		Data quality	User reviews	Reviews by marketplace				
	Customer relationship	Contract	Negotiated	Standardized				
Value delivery	Channels	Platform access	Closed	Open				
	Key resources	Platform infrastructure	Centralized	Decentralized				
	Key activities	Data processing activities	All	Limited				
Value capture	Revenue streams	Revenue streams	Usage based	Package pricing	Flat fee tariff	freemium	No info	
	Pricing model	Data pricing mechanism	Set by data marketplace owner	Set by sellers	Set by buyers	Auction	Negotiation	Real-time market
		Payment currency	Crypto	Fiat				

Figure 21. This taxonomy serves as the adapted version of the already developed taxonomies from Fruhwirth et al. (2020) and Spiekermann (2019). Dimensions are copied, adapted and excluded as explained in 4.4. The taxonomy is further refined by applying the empirical-to-conceptual approach and induce dimensions and characteristics from classifying existing data marketplaces.

	Component	**Dimension**	**Characteristic**					
Value creation	Customer segment	Domain	To be defined					
		Participants	To be defined					
	Value proposition	Privacy	Anonymized	Encrypted				
		Data source	Government	Social media	Commercial	Self-generated	Community	
		Data output	Transformed data	Non-transformed data				
		Data quality	User reviews	Reviews by marketplace				
	Customer relationship	Contract	Negotiated	Standardized				
Value delivery	Channels	Platform access	Closed	Open				
	Key resources	Platform infrastructure	Centralized	Decentralized				
	Key activities	Data processing activities	All	Limited				
Value capture	Revenue streams	Revenue streams	Usage based	Package pricing	Flat fee tariff	freemium	No info	
	Pricing model	Data pricing mechanism	Set by data marketplace owner	Set by sellers	Set by buyers	Auction	Negotiation	Real-time market
		Payment currency	Crypto	Fiat				

Figure 21: Preliminary taxonomy

E. Content analysis

This appendix presents the content analysis of the documents on the data marketplaces from the online desk research. The documents are retrieved from 4 online sources (see Table 31). These are the website of the respective data marketplace, the terms and conditions, the whitepapers and relevant articles from an external source.

Table 31: Data sources

Name	**Source type**	**Source of document**	**Reference code**
TomTom	Website	https://www.tomtom.com/	TT-1a
		https://developer.tomtom.com/	TT-1b
	Terms and conditions	https://www.tomtom.com/en_gb/legal/terms-and-conditions/	TT-2a
		TomTom (2017, December). GENERAL TERMS AND CONDITIONS FOR THE PROCUREMENT OF PRODUCTS AND SERVICES	TT-2b
	Whitepaper	How TomTom's HD Traffic and IQ Routes data provides the very best routing	TT-3a
		TomTom (2019) Annual Reports & Accounts 2019	TT-3b
	Forbes	O'Marah, K. (2017, April 3). Near-Death Experience: TomTom Rises From The Grave.	TT-4a
		Phelan, D. (2020, February 11). Huawai Makes Surprise TomTom Deal To Dodge Trump's Ban	TT-4b
		Koksal, I. (2020, February 10). Amazon Adds Auto-Specific Alexa Skills To Further Impact the Car Industry	TT-4c
Inrix	Website	https://inrix.com/	IN-1
	Terms and conditions	https://inrix.com/site-terms/	IN-2
	Whitepaper	Inrix OpenCar: Connected and Integrated Services for a Safe & Optimized Driving Experience. A Frost & Sullivan White Paper.	IN-3
	Forbes	Newcomb, D. (2016, September 26). Inrix Expands Parking Serice to more Mercedes-Benz vehicles, Others on the Horizon	IN-4a
		Yvkoff, L. (2015, September 10). Inrix's Acquisition of ParkMe Could Be a Game-changer for Navigation	IN-4b
		Newcomb, D. (2015, November 3). Volvo Signs With Inrix To Supply Global Real-Time Traffic Information Service	IN-4c
		Bruner, J. (2010, April 1). Compute Your Way Through Traffic.	IN-4d
	Harvard Case Study	Applegate, L.M. & Johnson, R. (2012, September). Inrix.	IN-5
HERE	Website	https://www.here.com/	HE-1a
		https://developer.here.com/products/platform/marketplace	HE-1b
	Terms and conditions	https://legal.here.com/en-gb/terms	HE-2
	Whitepaper	HERE location services on premises white paper	HE-3a
		HERE navigation on-demand	HE-3b
	Forbes	Singh, S. (2015, Augustus 5). HERE Acquisition By The Germans: Open Innovation On The Cards	HE-4a
		Newcomb, D. (2016, June 27). Inside Audi, BMW and Daimler's $3 Billion Bet On HERE's Mapping Business.	HE-4b
Caruso	Website	https://www.caruso-dataplace.com/	CR-1
	Terms and conditions	https://www.caruso-dataplace.com/privacy-policy/	CR-2
	Whitepaper	Naab, M. & Knodel, J. (n.d.). Architecture of the Caruso Ecosystem.	CR-3
		Software Engineering Institute (2018, June 4). SATURN 2018 Talk: Architecture of the CARUSO Ecosystem, by Matthias Naab. Youtube. https://www.youtube.com/watch?v=UlqdBsh_MF4&t=1010s	CR-3

	Automat	Bounie, D., Marcocchia, G. & Quinn, M. (2018, April 5). Automotive Big Data Marketplace for Innovative Cross-sectoral Vehicle Data Services.	CR-4
IOTA	Website	https://www.iota.org/	IOTA-1
	Terms and conditions	https://www.iota.org/research/privacy-policy	IOTA-2
	Whitepaper	Popov, S. (2018, April 30). Tangle	IOTA-3
	Forbes	Ponciano, J. (2017, November 28). IOTA Foundation Launches Data Marketplace for 'Internet-Of-Things' Industry	IOTA-4a
		Munford, M. (2018, January 23). Volkswagen's Chief Digital Officer Joins Blockchain Non-Profit Iota Foundation	IOTA-4b
		Fenech, G. (2018, November 20). IOTA – Fulfilling the Promise of Blockchain.	IOTA-4c
Ocean Protocol	Website	https://oceanprotocol.com/	OP-1
	Terms and conditions	https://oceanprotocol.com/privacy/	OP-2
	Whitepaper	Ocean Protocol Foundation (2019, April 15). Ocean Protocol: A Decentralized Substrate for AI Data & Services Technical Whitepaper	OP-3a
		Ocean Protocol Foundation (2017, October 19). A decentralized data exchange protocol, powered by blockchain technology and a cryptotoken - Business Strategy	OP-3b
		Ocean Protocol Foundation (2019, February). A decentralized data exchange protocol to unlock data for artificial intelligence – Reference Marketplace Framework	OP-3c
	Forbes	Wolfson, R. (2018, November 20). Diversifying Data With Artificial Intelligence And Blockchain Technology.	OP-4a
		Corea, F. (2018, October 4). The Blockchain-Enabled Intelligent IoT Economy	OP-4b

E.1 TomTom Codes

Table 32 shows the codes which are derived from content analysis on the case documents of TomTom. The analysis is performed in Atlas.ti and the table includes some examples of the coded lines, the initial codes and the focused codes.

Table 32: Content analysis TomTom

Dimension	**Focused coding**	**Initial coding**	**Examples of coded lines**
Domain	Traffic	In-vehicle navigation / advanced driver-assistance / autonomous driving / traffic information	"HD Traffic real-time traffic information is the backbone of a time-dynamic navigation concept that guarantees reliable routing and precise travel time information" (TT-3a)
Participants	Automotive customers Enterprise customers Consumer customers	Car manufacturers / tier 1 suppliers / technology companies / geographical information systems providers / government bodies / traffic management institutions / Automotive customers / Enterprise customers / consumer customers	"Our Automotive business unit licenses to automotive customers – both automakers, known as original equipment manufacturers (OEMs), and head-unit vendors, known as Tier 1 suppliers." (TT-3b)
Platform access	Closed	User account	"You may be required to create an Account and subsequently to log into the relevant Platform in order to: i. order a Product or Service; ii. access or use a Service, App and/or User Contributed Data (either obtained directly via TomTom or via a Reseller), and iii. be able to upload and/or download User Data and/or User Contributed Data. iv. manage your account and/or subscriptions." (TT-2a)
Privacy	Anonymous	Use information for the purpose and duration for which it was obtained / anonymous GPS data	"The profiles have been compiled by aggregating 1.4 billion of anonymous GPS probe data, shared by TomTom's broad user community" (TT-3a)
Data source	Government, commercial, self-generated	TomTom source / user data / user contributed data / GPS data / GSM data / cellular floating phone data / incident context data / TMC 3rd party messages / historic speeds	"The core sources of traffic data collection systems are probe data from cell phone operators in the various countries as well as GPS probes from the installed base TomTom connected devices and commercial fleets with TomTom WORK navigation systems" (TT-3a)
Data output	Aggregated data	Maps / navigation software / traffic information / dynamic routing / shift to wearables / product is the core of an ecosystem / retail sell-through data / online maps	"Meanwhile, the data is an essential ingredient of TomTom Personal Navigation Devices and online maps, and is branded as IQ Routes." (TT-3a)
Data service	Customized location analysis	License products / offer tailor-made solutions / Monitor laws and regulations / need for diversification of content, telematics and licensing / payment service / deliver product / physical delivery / digital delivery / repair broken product / shift towards	"We license maps, navigation software and online services as components for applications, offering tailor-made solutions to meet customer's specific needs" (TT-1a)

		personalization / customized supply response / build mapping app with partner / supply devices to businesses	
Data quality	Reviews by marketplace	Check quality	"The clusters undergo a couple of quality and alignment steps, finally leading to a set of profiles which serves all countries and markets or a given version of the map" (TT-3a)
Contract	Bilateral	Bilateral contract / accept proposal	"Contracts are concluded only after TomTom has accepted the Supplier's (final) quotation, proposal, or offer by issuing a Purchase Order" (TT-2b)
Market positioning	Independent	Independent	-
Platform infrastructure	Centralized	Store the data	"On a road segment, only the reference speed (normally the free flow speed) and a link to the most similar profile per day of the week are stored, instead of all the speed data for each time per day." (TT-3a)
Data processing activities	All	passive channel for distribution of user contributed date / monitor user contributed data / enhance data with road data / enhance data with weather information / filter unnecessary data / separate data / analyze data / filter data / compare real-time data to historic data / predict data / collect own data / link data points / create speed profiles / provide dynamic routing / estimate arrival time / harmonize data / integrate companion platforms / integrate marketing strategies / integrate supporting systems /	"For filtering, an enhanced data analysis is necessary, for example to separate handsets which are used in a train. As a typical speed pattern appears when calls are coming from trains, because all handsets have the same speed and handover events, these data can be taken out." (TT-3a)
Key partners	Joint venture	Acquired Tele Atlas before crisis / partnership with Nike / partner Huawei for mapping needs / partner with Amazon for better navigation	"First came the global financial crisis, including the collapse of Lehman Brothers. That was bad enough for any business, but especially bad for TomTom, which just before the crash had paid three billion Euros for a mapping company called Tele Atlas and was now loaded with debt." (TT-4a)
Revenue streams	Usage based, flat fee tariff	Free / monthly subscription / single basis / subscription bases / trial subscription / no more income from Sat Nav box sale / end-user subscriptions / pay as you grow	"A Service may be offered on a single basis or on a subscription basis. When you have a paid subscription, the Service will be provided to you either in a monthly or a fixed basis. Monthly Paid Subscriptions will be entered into for an indefinite period of time, and you will be automatically charged to your preferred payment method at the start of each subscription period for the fees and taxes applicable, unless you cancel the monthly subscription BEFORE the subscription period is renewed. Subscriptions will be

			entered into for an indefinite period of time, unless TomTom and you have agreed on a fixed period of time." (TT-2a)
Data pricing mechanism	Set by data marketplace	Set by data marketplace	"All prices quoted by Supplier shall be fixed, on a time and material basis or as otherwise detailed or agreed to in the Specifications and/or the Purchase Order, expressed in euros (unless stated otherwise), without prejudice and subject to these General Terms and Conditions, exclusive of VAT but inclusive of any other taxes, incidental costs and/or expenses." (TT-2b)
Payment currency	Euro	Euro	-

E.2 INRIX Codes

Table 33 shows the codes which are derived from content analysis on the case documents of INRIX. The analysis is performed in Atlas.ti and the table includes some examples of the coded lines, the initial codes and the focused codes.

Table 33: Content analysis INRIX

Dimension	**Focused coding**	**Initial coding**	**Examples of coded lines**
Domain	Traffic	Automotive industry / transportation industry / real-time traffic information RTTI / traffic data	"Primarily traffic and connected services" (IN-5)
Participants	Internal developers	OEM / TIER 1 / app developers / fleet-based service providers / media / mobile broadband platform providers / public sector / website / automotive / media	"Our partnership with Volvo will enhance the luxury experience Volvo drivers around the globe have become accustomed to by providing them with real-time traffic data" (IN-4c)
Platform access	Closed	Closed	-
Privacy	anonymous	Data control and privacy decided by OEM / API compliance / anonymous data / anonymous real-time GPS data	"The GPS installed in the vehicles acted as a sensor for the RTTI firm's network, sending information such as car location, speed and direction of travel anonymously back to a central processing center every minute" (IN-5)
Data source	commercial, self-generated	Road sensors / vehicles / GPS-equipped vehicles / mobile phones / network-based cellular car data / netFCD / users actively report data / radar sensors in parking lots / logistics fleet customers / proprietary app users	"By 2012, its network included over 100 million vehicles and devices, including GPS-equipped commercial vehicles, consumer vehicles, mobile phones and data from road sensors, all of which fed INRIX with up-to-the minute data on location, speed and directional heading." (IN-5)
Data output	Aggregated data	Real-time parking / traffic information / voice-enabled technology / location data / speed data / direction data / state of traffic / average speed / alternative routes / estimated delay times / crowd-sourced data / create lock-in effects / deliver parking info in dashboard / provide data to navigation providers	"INRIX delivers innovative products for the automotive and transportation industries such as real-time parking and traffic information and solutions that facilitate the safe testing and deployment of autonomous vehicles." (IN-1)
Service	Customized location analysis	Differentiate branding of OEM / take away negotiation between OEM and developers / customized services / aid driver in fast arrival / solve congestion / customized traffic data / different pricing for different customers / selling data as a service / train employees to use tools / pre-paid parking lots / sell aggregated data back to manufacturers / supply predictive traffic flow information	"Support for automakers' custom applications" (IN-3)

Data quality	Reviews by the marketplace	Reviews by the marketplace	"INRIX combined this data with proprietary algorithms to produce much higher quality" (IN-5)
Contract	Bilateral	OEM contracts are long term / deep customer relationship / due to contract term	"Over the next three years our revenue is projected to grow substantially based on deals we already have signed because the nature of the automotive business is such that you sign a deal, then you have to wait two years until the cars start shipping and then you start getting revenue." (IN-5)
Market positioning	Independent	Independent	-
Platform infrastructure	Centralized	Cloud environment / central data storage	"A cloud environment creates the ability for app developers and trusted app brands to deliver up-to-the-minute contextual content due to its standardized functionality." (IN-3)
Data processing activities	All	Analyze data / interface smartphone / find location / get directions / identify nearby parking spots / send parking alert / send safety alert / send traffic alert / update content / connect written software from OEM / aggregate third-party content / reduce down-time / reduce time-to-market / apply proprietary algorithms / crowd-sourcing vehicle data / collect GPS data / algorithms analyze data / statistically estimate RTTI / invest in algorithms / invest in distribution / invest in processing / collect data / predict traffic data / map data / combine GPS data with other traffic affecting data / collect data over prolonged time period / make algorithms / process data / separate data / invest money / develop applications and services / embed transaction-based parking reservations / predict EV range based on traffic / compare current speed to speed limit / detect slow-downs	"Our idea was to collect data from them on how fast their vehicles were going and where they were. This data would be transmitted back to our servers. From there we could process the data with a set of algorithms, to create accurate understanding of traffic conditions. We could then transmit real-time traffic reports to our customers, selling the data as a service." (IN-5)
Key partners	Strategic: Distribution, software, Sales	Content integrators / license traffic prediction technology from Microsoft / partnership with fleet / partnership with Microsoft for data management technology / partnership with Tele Atlas for sales / Clear Channel data distribution / acquisition of ParkMe / partnership with OEM to provide traffic data	"The technology developed by Microsoft Research was a sophisticated set of algorithms capable of accounting for dozens and dozens of variables" (IN-5)
Revenue streams	Usage based, flat fee tariff	Per subscriber per month / licensing fee per mile covered per year / revenue comes after car shipping	"Pricing differed based on customer type typically relying on either subscriber fees or licensing fees." (IN-5).

Data pricing mechanism	Set by data marketplace	Set by data marketplace	-
Payment currency	Euro	Euro	-

E.3 HERE Codes

Table 34 shows the codes which are derived from content analysis on the case documents of HERE. The analysis is performed in Atlas.ti and the table includes some examples of the coded lines, the initial codes and the focused codes.

Table 34: Content analysis HERE

Dimension	Focused coding	Initial coding	Examples of coded lines
Domain	Traffic	cross-industry data / location-centric data / location information / maps / RTTI / Weather data	"In the HERE Open Location Platform, data consists of both maps and location information that HERE provides, such as Real-Time Traffic and Weather, as well as data that you and other users provide." (HE-1b)
Participants	Internal and external analysts	Integrate third-party services and OEM / maintain relationship between OEM and consumer / Data provider / data consumer / third parties / business analysts / business scientists / developers / government	"Another example is a federal government that needs to route its heavy vehicles along appropriate roads, taking into account legal and physical vehicle restrictions and complying with data protection and security rules." (HE-3a) "HERE Neutral Server functions as a single point of data access from the vehicle manufacturer, which they can use to enhance the driving experience for the customers." (HE-1b)
Platform access	Open	Request access to data provider / control information a customer can see / view listing on invitation	"The request process makes it simple for providers to manage the level access to their datasets. All requests are logged and recorded by the system." (HE-1b)
Privacy	Encryption	Private listing / public listing / semi-private listing / continuous access until deactivated / encrypted data	"Data stored at rest in versioned layers, stream layers and index layers is encrypted using AES-256, a strong, proven, block cipher." (HE-1b)
Data source		Sensor data / neutral server to distribute data / data provided by HERE / data from other users	"In the HERE Open Location Platform, data consists of both maps and location information that HERE provides, such as Real-Time Traffic and Weather, as well as data that you and other users provide." (HE-1b)
Data output	Aggregated & Standardized data	Catalog metadata / three-dimensional maps	"Today, we're creating living three-dimensional maps that grow upwards, breathing with layers of information and insights." (HE-1)
Service	Customized location analysis &	Ensure functionality with poor network coverage / Connected car navigation / on demand service / provide a HERE SDK / payment and billing partnership	"An example of a location-based application involving large amounts of location data and high volumes of service

	standardized data transfer	/ cloud APIs / data services / map software / UX elements / individual OEM branding / modularity lowers costs / Find data / subscribe to data / consent management / provide software development kit / provide input and output metrics / enable data promotion, sharing and licensing / facilitate interaction and data exchange / provide monitoring tools / catalog management permission / cooperate in teams / publish data to catalog / have an OLP account / advertise data in listing / convert currency / no support for license execution / filter data / search catalogs	transactions, is location-based analytics, applied to improve the user experience and to gain geospatial insights from big chunks of data." (HE-3a)
Data quality	No info	-	-
Contract	Bilateral & standardized	Subscription functions as contract / standard terms and conditions / self-accept contract / no human interaction / subscribe to listing / provide metadata / provide SDK / visualize data	"There are three types of subscription options available: Customized - This is based on terms that are negotiated offline. Evaluation - These terms and conditions are available for you to self-accept and subscribe to from the listing, as there is no subscription flow. This type of subscription allows you to read data with no subscription fee. Commercial - These terms and conditions help Providers create their listings with different pricings for Marketplace, and allow Consumers to sign up with no human interaction so that the Marketplace can generate usage and/or line-items on the Consumer's bill." (HE-1b)
Market positioning	No info	-	-
Platform infrastructure	Centralized	Three element architecture / edge technology/ connect different IT layers / access the HERE Cloud / minimum quality guaranteed by edge / modular architecture/ HERE Cloud for larger storage / maintain and update navigation services / Store data in catalogs / store data in partitions / generic partitioning / tile partitioning / limited data storage	"To ensure good performance, the HERE Open Location Platform has limits on data storage and throughput. Some limits can be controlled by layer configuration, which may impact your cost since you are charged based on how you have configured the layers and data usage." (HE-1b)
Data processing activities		Integrate data in the marketplace / configure consumer navigation system / combine OEM portal and HERE NavOD SDK / List data / aggregate real-time parking data / catalog data / report input and output data metrics / segment data layers / support consistent content / standardize data using schemas / provide schemas	"A catalog is a data storage, that is available for the Data Provider or Data Consumer to access at any time." (HE-1b)

Key partners	-	-	-
Revenue streams	Usage based	Different subscription / usage based / free / standard pricing	“Cloud-based solutions offer low entry costs and for many businesses, attractive on-demand charging based on actual usage.” (HE-3a)
Data pricing mechanism	Negotiation & set by seller & set by data marketplace	Negotiate contract offline / set by sellers / negotiate contract	“With a Commercial subscription process, customers can subscribe to the listing with default pricing by clicking Subscribe or Review pricing for the listing with customized pricing.” (HE-1b)
Payment currency	Euro	Convert currency	“Customers will pay the currency they’ve chosen for their HERE account. HERE will convert currency based on exchange rates at the time of billing.” (HE-1b)

E.4 Caruso Codes

Table 35 shows the codes which are derived from content analysis on the case documents of Caruso. The analysis is performed in Atlas.ti and the table includes some examples of the coded lines, the initial codes and the focused codes.

Table 35: Content analysis Caruso

Dimension	**Focused coding**	**Initial coding**	**Examples of coded lines**
Domain	Automotive	Smart automotive domains	"smart mobility" (CR-4)
Participants	External third parties	Retrofit suppliers (short-term) / OEM (long-term) / only one OEM / aftermarket / service providers / founding partners are data sellers	
Platform access	closed	-	-
Privacy	Anonymous	Anonymous data / processing under expressed consent / adhere to privacy laws / founders act as legal guarantee	"As a rule, the analyses of your browsing patterns are conducted anonymously; i.e. the browsing patterns cannot be traced back to you." (CR-2a)
Data source	OEM, Tier 1	Connected vehicle / connect other marketplaces as well / vehicle data	"Development of connected vehicles in Europe" (CR-4) "Founding partners provide the customer base for data suppliers" (CR-4)
Data output	Standardized data	Vehicle information / in-vehicle data / process data / metadata / master data / configuration data / reference data / transactional data / streaming data / monitoring data	"transactional data | streaming data | monitoring/logging data | rather dynamic, high volume" (CR-4)
Data service	Data brokering service	Service marketplace / align business models of data sellers, buyers and marketplace / support client about thinking what data to upload and when to join marketplace / account management service / billing service / file upload service / statistics service / marketplace service (index, offer, subscription) / connecting correct stakeholders / influence partners in ecosystem to make trade happen / provide access to datasets	"We are focusing on providing the data brokering infrastructure and are not offering data or services ourselves." (CR-1)
Data quality	No info	-	-
Contract	Bilateral	Contract with data marketplace / contracts between provider and consumer	"Whether you want to consume or provide data – or do both, like the majority of our partners – it all starts with our partner agreement." (CR-1)
Market positioning	Sellers	Develop our platform according to the needs of our partners / owned by participants / website operator processes data / private / 34 shareholders	"We are expanding our shareholder base constantly across other segments of the mobility market bringing us even closer to achieving our goal of creating the mobility data & service marketplace owned by its participants." (CR-1)

Platform infrastructure	Central	Ecosystem involves many parties and layers / central Caruso platform / marketplace layer (meta-data) / brokering engine layer (API) / Amazon Web Services platform	"Providing a secure platform and respecting the privacy of data owners is vital to the future success of the whole ecosystem. As a result, our platform is secure by design, built on AWS with Auth0 authentication, plus every data and service that a company wants to onboard is checked for GDPR-compliance." (CR-1)
Data processing activities	Limited	No data offering / no services / provide brokerage infrastructure / govern the marketplace / analyze use patterns / collect / harmonize / distribute / match buyers and sellers	"All the data available on the platform is being harmonised to match into a precise data index, providing all partners full transparency of the exact data points available." (CR-1)
Key partners	-	Stakeholder verticals	"We are growing our investors in a structured manner with a balanced portfolio from all our stakeholder verticals to provide long-term sustainability for all of our partners." (CR-1)
Revenue streams	Margin	Marketplace gets margin / can be any revenue model / revenue for brokering, contracts handling and data analysis	"Caruso revenue scheme appears to be based on broker transactions for B2B data exchange and in services for contracts handling and data analytics." (CR-4)
Data pricing mechanism	Set by seller	Control over data and prices at data seller	"You have full control of your data & service and prices." (CR-3)
Payment currency	Euro	PayPal payment	"If you choose payment via PayPal, we will share the information you enter with PayPal (CR-2)

E.5 IOTA Codes

Table 36 shows the codes which are derived from content analysis on the case documents of IOTA. The analysis is performed in Atlas.ti and the table includes some examples of the coded lines, the initial codes and the focused codes.

Table 36: Content analysis IOTA

Dimension	Focused coding	Initial coding	Examples of coded lines
Domain	All industries	Interconnected and autonomous devices / machine economy / IoT industry / fine granular datasets / multiple data domains	“The marketplace is hinged on data's use applications in industries such as supply chain, smart cities, energy, manufacturing and healthcare.” (IOTA-4a)
Participants		Participants who issue transactions / participants who approve transactions / corporations / allow B2B, researchers and hobbyists / do not know who or how it will be used / any industry who have transactions / Governments / industries / users and validators are the same	“There are two distinct types of participants in the system, those who issue transactions, and those who approve transactions.” (IOTA-3) “The beauty of enabling fine-granular trade access is that we really do not know who or how it will be used, except that we know it is a completely new paradigm.” (IOTA-4a) “The way this is resolved is that instead of having these two parties turn the users into validators, we make the users and the validators one and the same.” (IOTA-4c)
Platform access	Open	Public marketplace / open data from sensors / open access	“The public marketplace aims to give connected devices the ability to securely transfer, buy and sell fine-granular and diverse datasets while ultimately facilitating access to data that oftentimes sits unused.” (IOTA-4a)
Privacy	Encrypted	Encrypted	-
Data source	Government, commercial, self-generated, community	Content published by the IOTA / corporates / governments	“Within the next few days, participants in these industries and others are expected to open and enable access to streams of data generated by sensors they've deployed.” (IOTA-4a)
Data service	Brokering service	Do research / develop software for the ecosystem / educate people about the foundation / standardize the economy of things / discuss applications with customers / promise of immutability / securing data for free	“Educate and promote technologies and use cases for new generations to understand and to ensure the Foundation’s success (IOTA-1a) “Standardize and ensure the maturity and widespread adoption of the economy of things” (IOTA-1a)
Data output	Standardized data	Solves scalability issue / data integrity	“When you put data onto the ledger, it has data integrity, which means it can never be changed again.” (IOTA-4c)

Data quality	User reviews	User verifies previous transactions / reward / validate two previous transactions / more activity = more validation / users must approve transactions / contribution to the network security / transaction not approved if transaction is in conflict / transaction request needs to approve previous transactions / level of confidence / nodes are not required to have consensus in validity / weight node proportional to invested work	"In order to make a transaction in the Tangle, two previous transactions must be validated with the reward for doing so being the validation of your own transaction by some subsequent transaction." (IOTA-1b)
Contract	Standardized	Trust in DLT for inclusive and permissionless economy / no need for trusted third-parties	"It's this decentralized permissionless ledger, where the data will be hosted, that will ensure the data being sold on IOTA's marketplace is tamper-proof." (IOTA-4a)
Market positioning	-	-	-
Platform infrastructure	Decentral	Open-source / distributed ledger / tangle individual transactions / directed acyclic graph / ledge stores transactions / distinguish between low load and high load / more decentral than blockchain	"The transactions issued by nodes constitute the site set of the tangle graph, which is the ledger for storing transactions." (IOTA-3)
Data processing activities	All	Synchronize ledgers of data and money / user performs computational work / design transaction rule for linking transactions / algorithm appoints linking transactions / check whether transactions conflict / motivate node to transact / drop lazy node / incentive to participate	"Every node calculates some statistics, one of which is how many new transactions are received from a neighbor. If one particular node is "too lazy", it will be dropped by its neighbors. Therefore, even if a node does not issue transactions, and hence has no direct incentive to share new transactions that approve its own transaction, it still has incentive to participate." (IOTA-3)
Key partners	-	-	-
Revenue streams	Usage-based Funding	Free to use / funded by donations / fee free / genesis transaction / initial balance of tokens / tokens distributed over founder addresses / non-profit / no bit for priority / digital payment without fee / pay for the exact quantity you use	"With IOTA, when you no longer have these fees, you can create a completely new economy where you pay for the exact quantity you use, rather than pay upfront or after the factor based on some statistical projections of how much you are going to use it." (IOTA-4c) "The IOTA Foundation is funded in three ways:1) Holdings of IOTA tokens from community donations and unclaimed tokens from the initial crowd sale. 2) Grants from governments to perform research and development. 3) Donations from individuals or enterprises." (IOTA-1)

Data pricing mechanism	Set by seller	Defined by sensor owner	-
Payment currency	Cryptocurrency	Tangle / cryptocurrency / IOTA cryptocurrency	“This approach is currently being implemented as a cryptocurrency called iota” (IOTA-3)

E.6 Ocean Protocol Codes

Table 37 shows the codes which are derived from content analysis on the case documents of Ocean Protocol. The analysis is performed in Atlas.ti and the table includes some examples of the coded lines, the initial codes and the focused codes.

Table 37: Content analysis Ocean Protocol

Dimension	**Focused coding**	**Initial coding**	**Examples of coded lines**
Domain	Full ecosystem	Autonomous vehicles / medical research	-
Participants		AI researchers / data scientists / data buyer / data curator / data marketplace / data seller / data verifier / data keeper / community / data consumers / data providers / developers / marketplaces / AI developers provide code as a service / lower barrier for smaller players to compete with tech giants	"The greatest beneficiaries are companies that have both vast data and internal AI expertise, like Google and Facebook. In contrast, AI startups have amazing algorithms but are starving for data; and typical enterprises are drowning in data but have less AI expertise. The power of both data and AI—and therefore society—is in the hands of few." (OP-3a)
Platform access	Open access	Decentralized access control / compute brought to the data / granting access permission / open access / control at data provider / blockchain monitors data access / blockchain provides control points	"while providing open access for developers to build services" (OP-3b) "The marketplaces built on Ocean Protocol will allow data to be accessed by all participants, ensuring that no central player can control or exploit the data." (OP-4a)
Privacy	Encrypted	Encrypted / keep control and privacy	"During the publishing process, the publisher provides the file URLs as plaintext, which will be encrypted by metadata store (Aquarius) in the backend and stored as encrypted URLs." (OP-3a)
Data source	Government, social media, commercial, community	Companies with data and AI expertise / enterprises with data without AI expertise / start-ups with algorithms but without data	"The greatest beneficiaries are companies that have both vast data and internal AI expertise, like Google and Facebook. In contrast, AI startups have amazing algorithms but are starving for data; and typical enterprises are drowning in data but have less AI expertise." (OP-3a)
Data service	Data brokering service	Providing licensing framework with pricing / incentivize data sharing / provide tools for discovery and value-added services / build open software / activate community of developers / provide pricing schemes / token storage / integrate metadata storage / integrate tools / provide dashboards / data compliance frameworks	"Support the designed token dynamics, including token storage and smart contracts business logic. Support for free, non-fungible, fungible and programmable pricing schemes." (OP-3b)
Data output	Standardized data	Normalize data / metadata content	"Ocean Protocol's processing functionality provides data

			curators with the ability to normalize exposed data in order to create new assets, while keeping track of source or background IP" (OP-3c) "OEP8 specifies the common attributes that must be included in any Asset Metadata stored in the Ocean Network, such as name, dateCreated, author, license, price, files (to URLs), file checksums, tags, and more. In addition, OEP8 recommends some additional attributes for discoverability and normalizes these attributes for curation purpose, which serve a common structure for sorting and filtering on DDOs." (OP-3a)
Data quality	Reviews by marketplace	Signal quality, reputation and ward / proof correct data file available / earn ocean tokens by providing data / data keeper / provide network rewards	One only gets network rewards for data they've staked if they also make it available when requested; making data available is a key role of keepers" (OP-3a)
Contract	Standardized	Verifiable service agreements / decentralized service agreements / smart contract	"First, a consumer (via Squid) conducts search on a marketplace's metadata store (via Aquarius interface). He/she finds a service offering (SEA) for data or compute that she likes. She digitally signs the SEA. In the next few steps, a service provider running Brizo will execute the agreement so that consumer can access (via on-chain access control) and consume the asset after sending the payment to Keeper smart contract." (OP-3a)
Market positioning	Independent	-	-
Platform infrastructure	Decentral	Decentralized orchestration / blockchain technology	"Ocean does decentralized orchestration: at its core are decentralized service agreements and decentralized access control, which execute on decentralized virtual machines." (OP-3a)
Data processing activities	Limited	Storing and promoting meta data, linking assets and services / connect to data / data curation / monetize data / connect services / inter-service network / metadata management / metadata storage / filter data assets / publish data assets / register users / search data assets / data standardization / provide computation, algorithms, storage, data / ocean keepers discover and validate / P2P consensus of ocean keepers / ocean	"Asset metadata is one part of Ocean DDOs. This is a JSON object with information about the asset." (OP-3a)

		verifier challenge data / cluster service providers in tribes	
Key partners	Contracting Infrastructure	Dutch Stichting for contracts / partnership with blockchain teams / partners to track compliance laws	"The Stichting lays out contracts for each node operator, including data processing agreements. This clarifies liability." (OP-3a)
Revenue streams	Any	Free / priced fungible / priced non-fungible / auction	"Ocean Protocol enables many types of data pricing strategies." (OP-3a)
Data pricing mechanism	Set by data seller	Pricing set by data seller	"set pricing for data via the Protocol to prevent vendor lock-in." (OP-3b)
Payment currency	Cryptocurrency	Tokens	"The Ocean Tokens (Ocean) are used on the Ocean Protocol network as the means of value exchange, to power the protocol and incentivize the keeper nodes of the network." (OP-3b)

E.7 Aggregated Codes per Dimension

The information in Table 38 shows the same codes as presented in table Table 32-Table 37, but is aggregated per business model dimension. Table 38 includes the codes from the dimensions which are included in the refined business model taxonomy, after iteration 7 during which all the codes of the data marketplace documents are reviewed.

Table 38: Aggregated codes

Domain	TT	In-vehicle navigation / advanced driver-assistance / autonomous driving
	IN	Automotive industry / transportation industry / real-time traffic information RTTI
	HE	cross-industry data / location-centric data / location information
	CR	Smart automotive domains
	IOTA	Interconnected and autonomous devices / machine economy / IoT industry / fine granular datasets / multiple data domains
	OP	Autonomous vehicles / medical research
Participants	TT	Car manufacturers / tier 1 suppliers / technology companies / geographical information systems providers / government bodies / traffic management institutions
	IN	OEM / TIER 1 / app developers / fleet-based service providers / media / mobile broadband platform providers / public sector / website / automotive
	HE	Integrate third-party services and OEM / maintain relationship between OEM and consumer / Data provider / data consumer / third parties / business analysts / business scientists / developers / government
	CR	Retrofit suppliers (short-term) / OEM (long-term) / only one OEM / aftermarket / service providers / founding partners are data sellers
	IOTA	Participants who issue transactions / participants who approve transactions / corporations / allow B2B, researchers and hobbyists / do not know who or how it will be used / any industry who have transactions
	OP	AI researchers / data scientists / data buyer / data curator / data marketplace / data seller / data verifier / data keeper / community / data consumers / data providers / developers / marketplaces / AI developers provide code as a service / lower barrier for smaller players to compete with tech giants
Data input	TT	TomTom source / user data / user contributed data / GPS data / GSM data / cellular floating phone data / incident context data / TMC 3rd party messages / historic speeds
	IN	Road sensors / vehicles / GPS-equipped vehicles / mobile phones / network-based cellular car data / netFCD / users actively report data / radar sensors in parking lots / logistics fleet customers / proprietary app users

	HE	Sensor data / neutral server to distribute data / data provided by HERE / data from other users / maps / RTTI / Weather data
	CR	Connected vehicle / connect other marketplaces as well / vehicle data
	IOTA	Content published by the IOTA / corporates / governments
	OP	Companies with data and AI expertise / enterprises with data without AI expertise / start-ups with algorithms but without data
Data output	TT	Maps / navigation software / traffic information / dynamic routing / need for diversification of content, telematics and licensing / shift to wearables / product is the core of an ecosystem / retail sell-through data
	IN	Real-time parking / traffic information / voice-enabled technology / location data / speed data / direction data / state of traffic / average speed / alternative routes / estimated delay times / crowd-sourced data / create lock-in effects / deliver parking info in dashboard / provide data to navigation providers
	HE	-
	CR	Vehicle information / in-vehicle data / process data / metadata / master data / configuration data / reference data / transactional data / streaming data / monitoring data
	IOTA	Solves scalability issue / data integrity
	OP	Normalize data
Data service	TT	Monitor laws and regulations / payment service / deliver product / physical delivery / digital delivery / repair broken product / shift towards personalization / customized supply response / build mapping app with partner / supply devices to businesses
	IN	Differentiate branding of OEM / take away negotiation between OEM and developers / customized services / aid driver in fast arrival / solve congestion / customized traffic data / different pricing for different customers / selling data as a service / train employees to use tools / pre-paid parking lots / sell aggregated data back to manufacturers / supply predictive traffic flow information
	HE	Ensure functionality with poor network coverage / Connected car navigation / on demand service / provide a HERE SDK / payment and billing partnership / cloud APIs / data services / map software / UX elements / individual OEM branding / modularity lowers costs / Find data / subscribe to data / consent management / provide software development kit / provide input and output metrics / enable data promotion, sharing and licensing / facilitate interaction and data exchange / provide monitoring tools / catalog management permission / cooperate in teams / publish data to catalog / have an OLP account / advertise data in listing / convert currency / no support for license execution / filter data / search catalogs
	CR	Service marketplace / align business models of data sellers, buyers and marketplace / support client about thinking what data to upload and when to join marketplace / account management service / billing service / file upload service / statistics service / marketplace service (index, offer, subscription) / connecting correct stakeholders / influence partners in ecosystem to make trade happen
	IOTA	Do research / develop software for the ecosystem / educate people about the foundation / standardize the economy of things / discuss applications with customers / promise of immutability / securing data for free
	OP	Providing licensing framework with pricing / incentivize data sharing / provide tools for discovery and value-added services / build open software / activate community of developers / provide pricing schemes / token storage / integrate metadata storage / integrate tools / provide dashboards / data compliance frameworks
Data quality	TT	Check data quality
	IN	Reviews by the marketplace
	HE	-
	CR	-
	IOTA	User verifies previous transactions / reward / validate two previous transactions / more activity = more validation / users must approve transactions / contribution to the network security / transaction not approved if transaction is in conflict / transaction request needs to approve previous transactions / level of confidence /

		nodes are not required to have consensus in validity / weight node proportional to invested work
	OP	Signal quality, reputation and ward / proof correct data file available / earn ocean tokens by providing data / data keeper / provide network rewards
Privacy	TT	Use information for the purpose and duration for which it was obtained / anonymous GPS data
	IN	Data control and privacy decided by OEM / API compliance / anonymous data / anonymous real-time GPS data
	HE	Private listing / public listing / semi-private listing / continuous access until deactivated / encrypted data
	CR	Anonymous data / processing under expressed consent / adhere to privacy laws / founders act as legal guarantee
	IOTA	Encrypted
	OP	Encrypted / keep control and privacy
Contract	TT	Bilateral contract
	IN	OEM contracts are long term / deep customer relationship / due to contract term
	HE	Subscription functions as contract / standard terms and conditions / self-accept contract / no human interaction / subscribe to listing / provide metadata / provide SDK / visualize data
	CR	Contract with data marketplace / contracts between provider and consumer
	IOTA	Trust in DLT for inclusive and permissionless economy / no need for trusted third-parties
	OP	Verifiable service agreements / decentralized service agreements / smart contract
Platform access	TT	User account
	IN	Closed
	HE	Request access to data provider / control information a customer can see / view listing on invitation
	CR	Closed
	IOTA	Public marketplace / open data from sensors / open access
	OP	Decentralized orchestration / blockchain technology
Platform infrastructure	TT	Store the data
	IN	Cloud environment / central data storage
	HE	Three element architecture / edge technology/ connect different IT layers / access the HERE Cloud / minimum quality guaranteed by edge / modular architecture/ HERE Cloud for larger storage / maintain and update navigation services / Store data in catalogs / store data in partitions / generic partitioning / tile partitioning / limited data storage
	CR	Ecosystem involves many parties and layers / central Caruso platform / marketplace layer (meta-data) / brokering engine layer (API) / Amazon Web Services platform
	IOTA	Open-source / distributed ledger / tangle individual transactions / directed acyclic graph / ledge stores transactions / distinguish between low load and high load / more decentral than blockchain
	OP	Decentralized orchestration / blockchain technology
Data processing activities	TT	passive channel for distribution of user contributed date / monitor user contributed data / enhance data with road data / enhance data with weather information / filter unnecessary data / separate data / analyze data / filter data / compare real-time data to historic data / predict data / collect own data / link data points / create speed profiles / provide dynamic routing / estimate arrival time / harmonize data / integrate companion platforms / integrate marketing strategies / integrate supporting systems /
	IN	Analyze data / interface smartphone / find location / get directions / identify nearby parking spots / send parking alert / send safety alert / send traffic alert / update content / connect written software from OEM / aggregate third-party content / reduce down-time / reduce time-to-market / apply proprietary algorithms / crowd-sourcing vehicle data / collect GPS data / algorithms analyze data / statistically estimate RTTI / invest in algorithms / invest in distribution / invest in processing / collect data / predict traffic data / map data / combine GPS

		data with other traffic affecting data / collect data over prolonged time period / make algorithms / process data / separate data / invest money / develop applications and services / embed transaction-based parking reservations / predict EV range based on traffic / compare current speed to speed limit / detect slow-downs
	HE	Integrate data in the marketplace / configure consumer navigation system / combine OEM portal and HERE NavOD SDK / List data / aggregate real-time parking data / catalog data / report input and output data metrics / segment data layers / support consistent content / standardize data using schemas / provide schemas
	CR	No data offering / no services / provide brokerage infrastructure / govern the marketplace / analyze use patterns / collect / harmonize / distribute / match buyers and sellers / provide access to datasets
	IOTA	Synchronize ledgers of data and money / user performs computational work / design transaction rule for linking transactions / algorithm appoints linking transactions / check whether transactions conflict / motivate node to transact / drop lazy node / incentive to participate
	OP	Storing and promoting meta data, linking assets and services / connect to data / data curation / monetize data / connect services / inter-service network / metadata management / metadata storage / filter data assets / publish data assets / register users / search data assets / data standardization / provide computation, algorithms, storage, data / ocean keepers discover and validate / P2P consensus of ocean keepers / ocean verifier challenge data / cluster service providers in tribes
Revenue streams	TT	Free / monthly subscription / single basis / subscription bases / trial subscription / no more income from Sat Nav box sale / end-user subscriptions / pay as you grow
	IN	Per subscriber per month / licensing fee per mile covered per year / revenue comes after car shipping
	HE	Different subscription / usage based / free / standard pricing
	CR	Marketplace gets margin / can be any revenue model / revenue for brokering, contracts handling and data analysis
	IOTA	Free to use / funded by donations / fee free / genesis transaction / initial balance of tokens / tokens distributed over founder addresses / non-profit / no bit for priority / digital payment without fee / pay for the exact quantity you use
	OP	Free / priced fungible / priced non-fungible / auction
Data pricing mechanism	TT	Set by data marketplace
	IN	Fixed
	HE	Negotiate contract offline / set by sellers / negotiate contract
	CR	Control over data and prices at data seller
	IOTA	Defined by sensor owner
	OP	Pricing set by data seller
Payment currency	TT	Euro
	IN	Euro
	HE	Convert currency
	CR	PayPal payment
	IOTA	Tangle / cryptocurrency / IOTA cryptocurrency
	OP	Tokens

F. Taxonomy Development Iterations

The taxonomy development iterations show the new or altered business model dimensions. Iteration 1-3 are performed to generate the preliminary taxonomy. Iteration 4-7 are performed after which the refined taxonomy is created. In iterations 4, 5 and 6 the characteristics of business model dimensions changed. Inclusion of this information in the development process would results in an information overload. Therefore, Table 39 only shows the development of the dimensions. The blue color stands for newly added dimensions and when a dimension is colored grey, it is removed.

Table 39: Taxonomy iterations

Step	*Iteration 1* Deductive	*Iteration 2* Inductive	*Iteration 3* Deductive	*Iteration 4* Inductive	*Iteration 5* Inductive	*Iteration 6* Inductive	*Iteration 7* Inductive
Value Creation	Customer segment	Customer segment	Domain	Domain	Domain	Domain	Domain
			Participants	Participants	Participants	Participants	Participants
			Market access				
			Integration				
	Value proposition	Value proposition	Pre-purchase testability				
			Access type				
			Privacy	Privacy	Privacy	Privacy	Privacy
			Data source	Data origin	Data input	Data input	
			Data output	Data output	Data output	Data output	Data output
				Data service	Data service	Data service	Data service
			Time relevancy				
			Data quality	Data quality	Data quality	Data quality	Data quality
	Customer relationship	Contract	Contract	Contract	Contract	Contract	Contract
Value Delivery	Channels	Channels	Platform access	Platform access	Platform access	Platform access	Platform access
	Key resources	Platform infrastructure	Platform infrastructure	Platform infrastructure	Platform infrastructure	Platform infrastructure	Platform infrastructure
	Key activities	Data processing activities	Data processing activities	Data processing activities	Data processing activities	Data processing activities	Data processing activities
	Key partners	Key partners	Key partners	Partnerships	Partnerships	Partnerships	
Value Capture	Revenue streams	Revenue streams	Revenue streams	Revenue streams	Revenue streams	Revenue streams	Revenue streams
	Pricing model	Pricing model	Data pricing mechanism	Data pricing mechanism	Data pricing mechanism	Data pricing mechanism	Data pricing mechanism
			Payment currency	Payment currency	Payment currency	Payment currency	Payment currency
	Cost model	Cost model					

New dimension

Removed dimension

G. Consent Form for Evaluation Interviews

This Appendix includes the consent form that is sent to the experts for the semi-structured evaluation interviews.

Consent Form for:

Data Marketplace Business Models

Please tick the appropriate boxes	Yes	No
Taking part in the study		
I have read and understood the study information dated __________, or it has been read to me.	□	□
I have been able to ask questions about the study and my questions have been answered to my satisfaction.	□	□
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	□	□
I understand that taking part in the study involves participation in an audio-recorded interview that will be transcribed as text. The audio recording will be destroyed after transcribing	□	□
Use of the information in the study		
I understand that information I provide will be used for the researcher's master thesis, (possibly) a scientific research article and educational purposes	□	□
I agree that my information can be quoted in research outputs	□	□
I agree that my real name can be used for quotes	□	□

Signatures

______________________	______________________	________
Name of participant	Signature	Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Rômy Bergman	___________________	________
Researcher name	Signature	Date

H. Evaluated Taxonomy

The refined taxonomy is evaluated with experts. They gave feedback on the taxonomy that helped improve the conciseness and comprehensiveness of the taxonomy. To improve the conciseness of the taxonomy, it is suggested to remove the data processing activities from the taxonomy. This dimension overlaps with the data service, in the value proposition. To improve the comprehensiveness of the taxonomy, the revenue streams characteristic *no info* is changed into *other*. This characteristic provides a loophole to classify data marketplaces that do not have the combinations of revenue streams as shown in the taxonomy. The adaptions result in the evaluated taxonomy, shown in Table 40.

Table 40: Evaluated taxonomy

<table>
<tr><th></th><th>Component</th><th>Dimension</th><th colspan="6">Characteristics</th></tr>
<tr><td rowspan="7">Value creation</td><td rowspan="2">Customer segment</td><td>Domain</td><td colspan="2">Location
(TT, IN, HE)</td><td colspan="2">Automotive
(CR)</td><td colspan="2">All industries
(IOTA, OP)</td></tr>
<tr><td>Participants</td><td colspan="3">Data sellers, data buyers, internal & external developers
(TT, IN, HE)</td><td colspan="3">Data sellers, data buyers & external developers
(CR, IOTA, OP)</td></tr>
<tr><td rowspan="4">Value proposition</td><td>Data service</td><td colspan="2">Customized map service
(TT, IN)</td><td colspan="2">Data brokering service
(CR, IOTA, OP)</td><td colspan="2">Both
(HE)</td></tr>
<tr><td>Data output</td><td colspan="2">Aggregated data
(TT, IN)</td><td colspan="2">Standardized data
(CR, IOTA, OP)</td><td colspan="2">Both
(HE)</td></tr>
<tr><td>Data quality</td><td colspan="2">Reviews by marketplace
(TT, IN)</td><td colspan="2">User reviews
(IOTA, OP)</td><td colspan="2">No info
(HE, CR)</td></tr>
<tr><td>Privacy</td><td colspan="3">Anonymized
(TT, IN, CR)</td><td colspan="3">Encrypted
(HE, IOTA, OP)</td></tr>
<tr><td>Customer relationship</td><td>Contract</td><td colspan="2">Negotiated
(TT, IN, CR)</td><td colspan="2">Standardized
(IOTA, OP)</td><td colspan="2">Both
(HE)</td></tr>
<tr><td rowspan="2">Value delivery</td><td>Key channels</td><td>Platform access</td><td colspan="3">Closed
(TT, IN, CR)</td><td colspan="3">Open
(HE, IOTA, OP)</td></tr>
<tr><td>Key resources</td><td>Platform infrastructure</td><td colspan="3">Centralized
(TT, IN, HE, CR)</td><td colspan="3">Decentralized
(IOTA, OP)</td></tr>
<tr><td rowspan="3">Value capture</td><td>Revenue streams</td><td>Revenue streams</td><td>Usage based
(TT, IN)</td><td>Usage based & freemium
(HE)</td><td colspan="2">Commission
(CR)</td><td>Donations
(IOTA)</td><td>Other
(OP)</td></tr>
<tr><td rowspan="2">Pricing model</td><td>Data pricing mechanism</td><td colspan="2">Set by data marketplace
(TT, IN)</td><td colspan="2">Set by seller
(CR, IOTA, OP)</td><td colspan="2">Both
(HE)</td></tr>
<tr><td>Payment currency</td><td colspan="3">Fiat currency
(TT, IN, HE, CR)</td><td colspan="3">Cryptocurrency
(IOTA, OP)</td></tr>
</table>

I. Ending Conditions

This appendix shows the objective and subjective ending conditions which should be met to design a complete taxonomy (Nickerson et al., 2013). Table 41 shows the application of the ending conditions for the business model taxonomy of data marketplaces. The satisfaction of the conditions is indicated with a cross. The iterations are represented per column. Iteration 7 is the revision of all data marketplaces for the respective dimensions and characteristics. During this iteration, no new dimensions or characteristics were added and no dimensions or characteristics were merged. Iteration 8 represents the evaluation interviews during which the subjective ending conditions were tested.

Table 41: Application of ending conditions

Ending Conditions		Iterations							
		1	2	3	4	5	6	7	8
OE1	All objects or a representative sample of objects have been examined							x	x
OE2	No object was merged with a similar object or split into multiple objects in the last iteration	x	x	x	x	x	x	x	x
OE3	At least one object is classified under every characteristic of every dimension					x	x	x	x
OE4	No new dimensions or characteristics were added in the last iteration							x	x
OE5	No dimensions or characteristics were merged or split in the last iteration	x	x	x		x	x	x	x
OE6	Every dimension is unique and not repeated (i.e., there is no dimension duplication)						x	x	x
OE7	Every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension)						x	x	x
SE1	Concise: the taxonomy is meaningful without being overwhelming								x
SE2	Robust: the dimensions and characteristics suffice to differentiate objects								x
SE3	Comprehensive: all objects can be classified								x
SE4	Extendible: new dimensions and characteristics can be added								x
SE5	Explanatory: the dimensions and characteristics explain an object								x