1 Introduction

"(…) there it is really necessary to collect the data and to interpret them and to process them to gain information and to create new business models"
„(…) da gilt es tatsächlich die Daten zu sammeln und die zu interpretieren und eben zu Informationen zu verarbeiten und daraus neue Geschäftsmodelle zu schnüren."

An interview partner answers the question whether data are the new oil in Logistics

1.1 Motivation

The Logistics sector is the third-largest industry branch in Germany, only ranking behind the Automotive and Commerce industries. It employs over three million employees spanning over 60.000, mostly small and medium-sized (SME) companies, and, as of 2019, generates revenue of about 279 billion euros (Bundesvereinigung Logistik 2018). These companies that usually transport goods from a start point to a destination via different modalities, e.g., by ship, truck, or train, face novel challenges posed by digitalization (Zanker, p. 10). According to a study by van Marwyk and Treppte (2016, p. 10), digitalization will drastically change the face of Logistics. The digitization process leads to the emergence of a considerably high amount and a wide range of logistical data. In the study, 36% of the respondents stated that digitization would partially change the Logistics industry, while 59% of the respondents stated that digitalization would completely change Logistics (see Figure 1).

Another study conducted by the *German Federal Ministry for Economical Affairs and Energy*[1] finds that the Logistics sector is generally underdeveloped in terms of digitalization. For example, the study finds that (BMWi 2018, pp. 13-22):

- The Logistics sector considers digitalization to be less important than the Service sector in general
- Highly digitized processes are less common in Logistics than in the Service sector in general

[1] In German: Bundesministerium für Wirtschaft und Energie (BMWi)

- Fewer Logistics firms make significant revenue with digital offerings than the Service sector in general
- Fewer Logistics firms consider the degree of digitalization in their service offerings as very high than in the Service sector in general

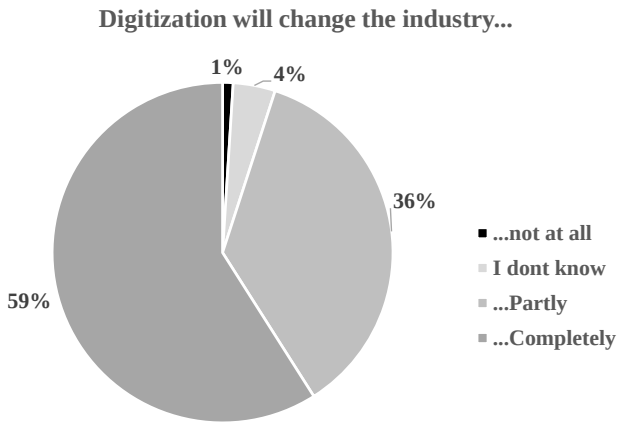

Figure 1: Results from a study in Logistics asking how digitalization will change the industry. Adapted from van Marwyk and Treppte (2016, p. 10)

More recently, a study by the BVL[2] finds that the Logistics industry currently undergoes a shift in digitalization perception (Kersten et al. 2020). While digitalization was often seen as a threat, there now is a trend showing companies recognizing its potential. The authors of the study suggest that this means that digital transformation needs to be proactively pushed forward (Kersten et al. 2020, p. 8). Compared to 2016, the study finds that, in 2020, most Logistics enterprises see the potential of digitalization and agree with the notion of digitalization being an opportunity for additional revenue generation and cost optimization (Kersten et al. 2020, p. 9). When asked about how business models are changed digitally, the study finds that over 50% of the respondents do not employ a digital business model or have a business line that focuses on it. The study even finds that 57% of the respondents either plan to digitize their business model in a time frame of either below or above five years. However, the largest segment of respondents indicated that there currently is no plan to transform the existing business model into a digital business model (see Figure 2).

[2] Bundesvereinigung Logistik E.V.

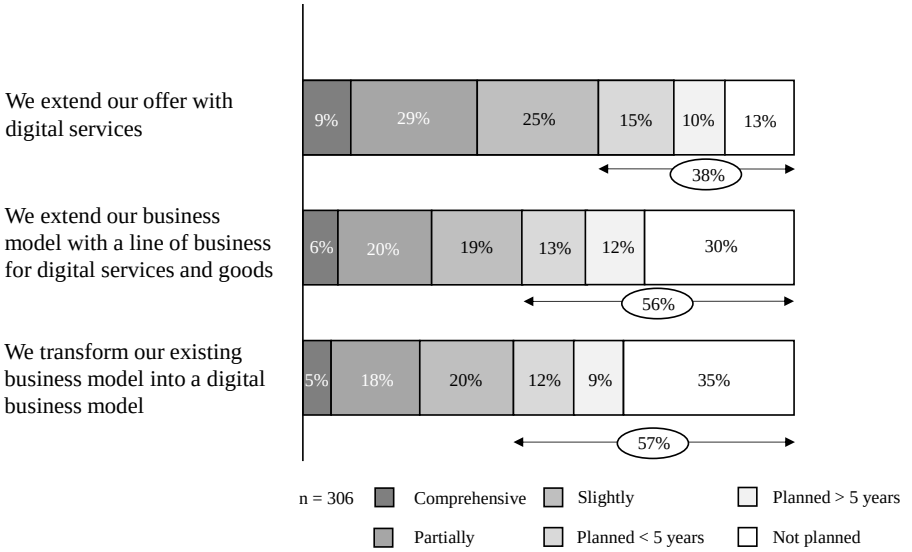

Figure 2: Results of a study asking how business models in Logistics are changed through digitalization (n=306). Adapted from Kersten et al. (2020, p. 16).

The study also finds that between the years 2020 and 2016, there has been a slight upward trend on all three questions regarding the maturity of digital business models (see Figure 3). The authors clearly identify the need for action to leverage the full potential of digital transformation for Logistics (Kersten et al. 2020, p. 17).

Summa summarum, what these studies show is an industry currently in a phase of disruptive and continuous shift, which is in the ongoing process of adjusting, adapting, and acclimating to digitalization. Thus, the studies clearly produce evidence that there is a need to assist practitioners in leveraging data for economic exploitation and to generate successful business models in Logistics. In that sense, they are the motivational bedrock for the thesis, as they indicate a practice-driven need for research on digital business models in Logistics.

That notion is supported by industry experts, that were interviewed during the thesis. Leveraging the treasure of data in Logistics is a commonly accepted paradigm that spurs opportunities for in-depth research. Correspondingly, the notion of the Logistics sector being traditional or sitting on a yet untapped treasure of data was a frequent point of discussion in the interview study outlined in Section 6. Illustratively, one industry professional described that issue as Logistics companies having to start comprehending the scope of the treasure of data they possess.

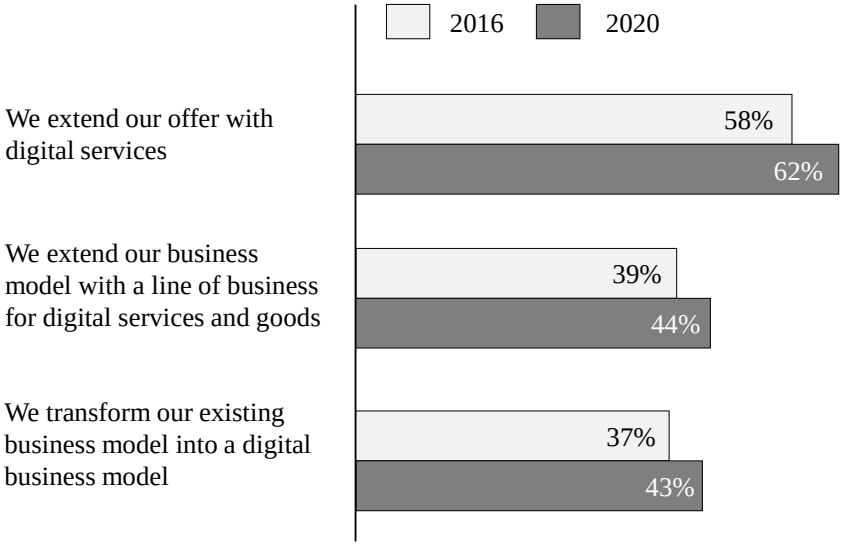

Figure 3: Comparison of answers regarding the state of digitalization of business models between 2016 (n=363) and 2020 (n=306). Adapted from Kersten et al. (2020, p. 17).

> *"That means, I believe that, in the Logistics sector, we only start to comprehend what kind of data treasure we have lying here. That means, currently, the understanding is not truly here, but if we go deep in us and think, there is an enormous treasure of data."*
>
> *"Das heißt, ich glaube in der Logistikbranche beginnen wir gerade erst zu begreifen, welchen Datenschatz wir da eigentlich liegen haben. Das heißt, aktuell ist das Verständnis noch gar nicht so wirklich da, wenn wir aber tief in uns gehen und uns das mal so überlegen, ist ein enormer Datenschatz vorhanden."*
>
> ***One interviewee is painting the picture of digitalization in Logistics.***

Thus, the need for systematic analysis and tools for assistance in handling and leveraging data for business model utilization is apparent and straightforward.

More recently, the emergence of logistical start-ups has led to the development of novel digital-borne business models that tackle and challenge long-established, traditional businesses and business models (CBInsights 2016; Oliver Wyman 2018, p. 5). The high number of newly emerging start-ups and the underlying considerable investment value indicates an evolving industry sector (Sucky and Asdecker 2019, p. 209). For example, contrary to traditional service delivery through controlling vehicles, some start-ups

provide marketplaces that mediate logistical services between enterprises or private drivers with spare capacity and those requiring transportation services. That phenomenon has since received growing attention both in academia (Göpfert and Seeßle 2017, pp. 1-2, 2018, pp. 1-2) and in managerial studies (Zimmermann, pp. 5-6) with the goal in mind to categorize, conceptualize, and align these start-ups. As one interviewee puts it, start-ups, contrary to incumbent firms, are especially suited to leverage data, as they are focused explicitly on that topic, i.e., they pursue a *green-field* approach free from legacy systems (Hartmann et al. 2016, p. 1383).

> *„But, especially if it is about collecting data, many start-ups are predestined because they really focus on the data topic."*
>
> *"Aber gerade wenn es darum geht jetzt viele Daten einzusammeln, sind viele Start-Ups prädestiniert, weil sie sich wirklich nur auf die Datenthematik erstmal geschmissen haben."*
>
> ***One interviewee highlighting the position of start-ups in digital business models.***

The findings generally support the notion that Logistics is relatively traditional, i.e., leaving many potentials to generate new offerings and prospects for optimization based on data (Heistermann et al., p. 37).

The thesis considers start-ups and incumbents as objects of investigation and uses the business model concept to analyze them (see Table 1). The idea of the *business model* is an essential analytical object in *Information Systems (IS)* research, as every business operates on an underlying operational logic, be it consciously or implicitly (Casadesus-Masanell and Ricart 2010, pp. 200-201; Teece 2010, p. 172). Subsequently, „All firms have business models (…)" (Gambardella and McGahan 2010, p. 262). That underlying logic explicates how the given business generates value for customers, delivers that value to the customers, and generates revenue from it (Teece 2018, p. 40). A business model is a tool relevant to strategic management decisions that exceed the mere consideration of profit formulas (Arend 2013, p. 395). In recent years, the business logic's conceptualization has become ever more prevalent and relevant in IS literature. Hence, the business model is a suitable tool to conceptualize business conduct and analyze start-ups and traditional enterprises. One can observe its utility by investigating the various publications producing *taxonomies*, *typologies*, or *classification schemes* with twofold objectives. Firstly, to decompose the structural composition of business models and, secondly, to establish comparability between them (e.g., see Timmers (1998)).

> *"The first compares scale models and role models to explain how the notion of business models enables us to classify businesses in a taxonomy or typology… valuable ways to expand our understanding of business phenomena and the development of ideal types."*
>
> ***Baden-Fuller and Morgan (2010, p. 157) explaining business models as models.***

Table 1: Exemplary differences between start-ups and incumbent firms.

Start-Ups	Incumbent Firms
Often follow high-risk and high-reward strategies, which also might lead to early failure (Cantamessa et al. 2018, p. 1)	Profit from prior innovations and control necessary assets (Criscuolo et al. 2012, p. 320)
Little established infrastructure, able to follow greenfield, digital strategies (Hartmann et al. 2016, p. 1383)	Have many legacy systems (Hartmann et al. 2016, p. 1383)
Often have a single business model (Sabatier et al. 2010, p. 443)	Often have multiple business models (Sabatier et al. 2010, p. 444)

Based on the pressure from leveraging data and the emergence of new start-ups that enter the domain of traditional Logistics, *logistics service providers* are faced with ideating new strategies and business models (Gruchmann et al. 2020, pp. 2-4). Evoking the famous notion of Darwin, only the fittest may survive, which in the present case means logistics service providers that adapt to the dynamic world of digitalization and that learn to navigate in it, leverage its potential and mitigate its risks.

> *"Further we must suppose that there is a power, represented by natural selection or the survival of the fittest, always intently watching each slight alteration in the transparent layers; and carefully preserving each which, under varied circumstances, in any way or in any degree, tends to produce a distincter image"*
>
> ***Darwin (2003, p. 174)'s famous quote explicating 'survival of the fittest'.***

1.2 Research Questions

This dissertation aims to clarify the role of data in digital business models in the Logistics domain and uncover their structure. It seeks to conceptualize the part of data as a key enabler for logistical business models.

The overarching research question is divided into five subordinate research questions. The frame of observation both covers *incumbent firms* and *start-ups*. As these types of businesses have drastically differing characteristics, they pose different research questions and require case-tailored research methods. For example, while start-ups often have an identifiable business model that can be conceptualized via *desk research*, incumbents usually have a portfolio of business models that are mostly not communicated transparently (see Section 4.2.2). Thus, contrary to desk research, gathering information from incumbents requires alternative modes of inquiry, e.g., interviews with experts from the domain as chosen in the present thesis.

The research questions follow the conventional practice, as they start very generally and broadly and get more specific and detailed alongside the research projects. Table 2 lists all research questions addressed in the dissertation and the corresponding sections and contributions.

Table 2: Summary of research questions and corresponding contributions.

#	Research Question	Contribution	Section
RQ	How are data used in digital business models in Logistics?	Specified for each RQ	-
RQ1	What are the central characteristics of digital business models in Logistics start-ups?	Taxonomy of digital business models in Logistics start-ups.	**4**
RQ2	What are the archetypes of digital business models in Logistics start-ups?	Archetypes of digital business models in Logistics start-ups	**4**
RQ3	What are the central characteristics and dimensions of data-driven business models in Logistics start-ups?	Taxonomy of data-driven business models in Logistics	**5**

RQ4	What is the context in terms of digitalization in German Logistics?	Explorative knowledge describing the context the interview study is embedded in.	**6**
RQ5	What are the design principles for data-driven business models?	Design principles for data-driven business models	**6**

Currently, a large variety of start-ups emerge that disrupt conventional Logistics. As of now, *AngelList*[3] (a database for start-ups) lists 2887 start-ups tagged as *logistics* enterprises. For example, these start-ups employ business models that range from *digital freight forwarders* mediating and organizing transportation services globally through digital booking platforms to those generating visibility in supply chains through dedicated tracking devices or mobile applications. The sheer number of start-ups and various business models make a clear identification of actual business models hard and motivate systematic inquiry. Thus, there is a need to structure that body of business models in Logistics start-ups, i.e., to "(…) structure or organize the body of knowledge that constitutes a field (…)" (Glass and Vessey 1995, p. 65). A suitable tool to achieve precisely that understanding of a domain of interest is to classify objects by generating a taxonomy (Szopinski et al. 2020, p. 2). Summarizing the above, the first research question reads as follows:

> ***Research Question 1 (RQ1):*** *What are the central characteristics and dimensions of digital business models in Logistics start-ups?*
>
> ***- Addressed in Section 4***

Inevitably, that taxonomy results in predominantly general characteristics and dimensions. Thus, to identify critical morphological configurations representing foundational and repeating business model patterns, the taxonomic findings are analyzed using *cluster analysis*. Through the generation of condensed clusters that contain similar objects (i.e., specific types of business models, such as transport marketplaces) and extraction of dominant configurations, one can derive *archetypical* patterns. These archetypical patterns help researchers and practitioners get to the heart

[3] https://angel.co/logistics last-accessed: 28.05.2020

of business model configurations in Logistics start-ups. Based on the findings generated in answering **RQ1**, the second research question reads as follows:

> ***Research Question 2 (RQ2):*** *What are the archetypes of digital business models in Logistics start-ups?*
>
> ***- Addressed in Section 4***

The third research question fluently builds on both **RQ1** and **RQ2**. In that, digital business models in Logistics start-ups reflect the nature of digital business models in general, as that they entail business models focussing either on value-generating mechanisms through leveraging *network effects* (platform business models) or *data as key resources* (data-driven business models) (Guggenberger et al. 2020a, pp. 7-9). The dissertation focuses on business models, explicitly leveraging data as the key resource, i.e., data-driven business models (see Section 3.2.2). As that sub-sample of business models is encompassed in the taxonomic analysis of **RQ1**, **RQ3** is answered using these results (as a specific taxonomy (McKelvey 1978, p. 1429)). Analyzing that sub-set in more detail with a more narrowly designed set of characteristics and dimensions, the third research question produces a specific taxonomy of data-driven business models and reads as follows:

> ***Research Question 3 (RQ3):*** *What are the central characteristics and dimensions of data-driven business models in Logistics start-ups?*
>
> ***- Addressed in Section 5***

Research questions four and five shift the frame of reference from start-ups to incumbent firms. Naturally, these pose different methodological challenges in acquiring data. As start-ups usually only have one business model and information about them are often quite transparently available, that is not the case with incumbent firms (Hartmann et al. 2016, p. 1383; Sabatier et al. 2010, pp. 442-443; Täuscher and Laudien 2017, p. 5287). These often have a wide variety of business models integrated into business model portfolios. Thus, one can not merely say that an incumbent Logistics business only does transport from A to B but instead has a host of various services (e.g., contract logistics or specialized logistical services). Additionally, contrary to being relatively small and demarcated, incumbent firms (as understood here) often have a

long-standing history and a large body of personnel and infrastructure. Thus, acquiring data is not easily done by *desk research* but requires tapping into the field through interviewing experts.

Adequate embedding and situating qualitative data collection requires a sound understanding of the studies' context. **RQ4** addresses that issue, as it generalizes findings from the qualitative interview study to transfer and demarcate a realistic image of the Logistics sector in Germany (as far as it is represented through the interviewees of the study). An initial indication about the Logistics sector and, correspondingly, its state of digitalization is given in Figure 2 and Figure 3. To gain insight into how experts of incumbent Logistics enterprises and applied research understand digitalization and digital/data-driven business models for their industry sector, the fourth research question is the following:

> ***Research Question 4 (RQ4):*** *What is the context in terms of digitalization in German Logistics?*
>
> ***- Addressed in Section 6***

Based on the data gathered in interviews, the last research question describes the transition from descriptive, analytical knowledge generated through taxonomies and descriptions of the state of Logistics to prescriptive knowledge. For that, the interview data are analyzed so that requirements for the successful design of data-driven business models in Logistics can be elicited. Based on requirements, the aim is to generate recommendations for action (i.e., design principles) to design successful data-driven business models in Logistics. Subsequently, the final research question is:

> ***Research Question 5 (RQ5):*** *What are the design principles for data-driven business models in Logistics?*
>
> ***- Addressed in Section 6***

Table 2 gives a summary of the five research questions and shortly explains the resulting contributions. Also, it points to the respective Sections in the thesis that explain each research question *en détail*. Figure 4, correspondingly, maps each research question to the suitable knowledge type that it produces, the frame of reference that it targets, and the analytical unit that it refers to.

1.3 Research Context

The present section outlines the research context that provides the basic framework of the thesis.

Digital Transformation: The context and enabling conceptual framework of this thesis is the ever-increasing *digital transformation* of industry sectors and social life. Illustratively, Kutzner et al. (2018, p. 1) evoke Charles Darwin's famous notion of "(…) survival of the fitness (…)" (Darwin 2003, p. 174) in evolutionary theory to paint a tangible analogy for firms needing to adapt in a new, digital environment, that is populated with complex, multi-faceted challenges (Al-Debei and Avison 2010, pp. 369-370). Digital transformation is the goal-oriented integration of digital technologies in organizations (Matt et al. 2015, pp. 339-340). Usually, one distinguishes between two underlying phenomena, coined *digitization* and *digitalization*, that may be defined as follows:

> "(…) the term *digitalization* has been coined to describe the manifold sociotechnical phenomena and processes of adopting and using these technologies in broader individual, organizational, and societal contexts."
>
> ***Legner et al. (2017, p. 301) defining digitalization.***
>
> "Digitization is the technical process of converting analog signals into a digital form, and ultimately into binary digits (…)"
>
> ***Legner et al. (2017, p. 301) defining digitization.***

Reis et al. (2018, pp. 417-418) provide a threefold classification of definitional approaches to *digital transformation* into *technological*, *organizational*, and *societal*. Most relevant for the present thesis, organizational definitions address the creation of new business models. For example, Hess et al. (2016, p. 124) define that digital transformation "(…) is concerned with the changes digital technologies can bring about in a company's business model, which results in changed products or organizational structure". In terms of business models, the frame of digital transformation entails a large variety of enablers (e.g., big data or digital technologies) that foster the development of novel products and services (Schallmo et al. 2017, pp. 5-6; Yoo et al. 2010, p. 732; Yoo et al. 2012, p. 1399), which trigger *digital innovation processes* (Dos Santos et al. 2014, p. 330) and require *digital business strategy* (Bharadwaj et al. 2013, p. 472).

Logistics: The thesis is embedded in the domain of Logistics. Logistics, in the thesis, means the transportation of goods from the point of origin to a destination or customer

(ten Hompel and Heidenblut 2011, p. 313). Moreso, the value-adding provision of logistics services, i.e., in terms of strategy and information (Topolšek et al. 2018, p. 1196). Logistics is especially suitable for data-driven innovation as the industry sector is, by its nature, striving for continuous optimization and has the potential to adopt new digital technology (Kayikci 2018, p. 784). For example, logistics service providers always want to choose the most optimal route with the explicit intention of saving costs. Also, B2C business models' services increase the pressure on B2B business models regarding the visibility of the transportation process (e.g., tracking & tracing).

Qualitative Research: The dissertation predominantly utilizes tools of qualitative research to generate data and produce findings. In qualitative studies, researchers endeavor to explain or interpret the phenomena of interest (Hunter 2005, p. 36). Meaning that the data require interpretation expressed through *judgment* (e.g., coding of interview or business data) (Saldaña 2015, p. 4; Sipe and Ghiso 2004, p. 482). The sources for qualitative data analysis range from *desk research* (see Section 4 and Section 5) on start-ups to interviews with incumbent firms (see Section 6). In the thesis, qualitative research includes coding of publically internet-based data of start-up firms (e.g., websites, blog articles, or YouTube videos) and textual transcripts of interviews with experts (Bowen 2009, pp. 27-28; Glaser 2007, pp. 49-50).

Information Systems Research: Information Systems research strives to produce relevant contributions to the knowledge base of information technology application in the organizational context (Hevner and March 2003, p. 111). The field diverges into two distinct research streams, namely *behavioral sciences* and *design sciences* (Alturki et al. 2012, p. 312; Hevner and March 2003, p. 111; March and Smith 1995, p. 253; Niederman and March 2012, pp. 2-3), although the validity of the dichotomous separation of both paradigms is the subject of discussion (Fettke 2008, pp. 45-46). Stahl (2009, p. 118) provides a condensed overview of both research strands shown in Table 3. The roots and foundations of each paradigm also diverge geographically, like the German-speaking countries (using the German term "Wirtschaftsinformatik") predominantly follow the design science research approach and the Anglo-Saxon sphere focusing more on the behavioral science research approach (Frank 2006, p. 1). The broadly generalized division is, however, subject to severe discussion amongst scientific peers in the IS community (see Baskerville et al. (2011) and Österle et al. (2011)).