10-11-2020

Advancing Data Monetization and the Creation of Data-based Business Models

Petri Parvinen
University of Helsinki, petri.parvinen@helsinki.fi

Advancing Data Monetization and the Creation of Data-based Business Models

Petri Parvinen

Strategic Marketing and Management
University of Helsinki
Industrial Engineering and Management
Aalto University
petri.parvinen@helsinki.fi

Essi Pöyry

Centre for Consumer Society Research
University of Helsinki

Robin Gustafsson

Industrial Engineering and Management
Aalto University

Miikka Laitila

Futurice

Matti Rossi

Information and Service Management
Aalto University

Abstract:

Although researchers have discussed big data for years, they have thus far has scarcely touched on directly selling and monetizing data assets. This aspect has particular relevance given recent concerns about data privacy and security and the simultaneous explosion in the use of data for marketing and service-development purposes. In this paper, we describe an empirical study on companies' initiatives about selling and monetizing data. We categorize the relevant business models based on several customer-refinement and scalability dimensions. We found several constraints (organization type, business type, data characteristics, privacy, and security) that companies should address to move from internally using data and supporting existing customers to generating new business through selling data. Based on the findings, we propose ways for practitioners to benefit from the data. For researchers, we provide directions for future studies that include developing strategies that foster compliance between companies' aspirations and consumer/societal restrictions and that facilitate data-based innovation and revenue generation.

Keywords: Data Monetization, Selling Data, Business Models, Constraints.

1   Introduction

The data that companies gather and store continues to increase exponentially. While researchers and practitioners have traditionally seen data as an internal asset that provides value to a firm's operations, they see it today more as an asset that can generate competitive value offerings to customers, added revenue streams to existing business, and even become a tradable asset (Chen, Chiang, & Storey, 2012; Choo, 1996; Delen & Demirkan, 2013; Opresnik & Taisch, 2015; Walker, 2014). We can already see the effects from firms that have used data in extended ways in their business and operations and undergone transformation as a consequence. Companies such as Uber, Airbnb, Facebook, Netflix, and Amazon have fundamentally changed their respective industries with data- and analytics-driven practices and offerings (Parmar, Mackenzie, Cohn, & Gann, 2014; Rossman, 2016; Spijker, 2014; Yousif, 2015). Innovating and designing data-based business models constitutes a key aspect in capturing value from data (de Reuver, Sorensen, & Basole, 2018).

Companies use data in many different ways. They may apply their improved data-based knowledge to develop their existing business in order to reduce costs or optimize production (internal purposes), or to create new data-based or data-informed market offerings (external purposes) (Choo, 1996). The literature to date has largely focused on internal usage and specifically on improving existing products and services typically in accordance with the concept of big data (Bean, 2017; Thomas & Leiponen, 2016). Since an obvious starting point for companies involves using big data internally—the so-called low-hanging fruit— many seek additional opportunities to use their data (Spijker, 2014). However, monetizing data by designing and selling novel data-based business offerings to external parties has attracted less attention in the academic literature. The few studies that do discuss the issue (Delen & Demirkan, 2013; Hartmann, Zaki, Feldmann, & Neely, 2016; Lewis & McKone, 2016; Najjar, & Kettinger, 2013; Wixom & Ross, 2017) operate at the conceptual or anecdotal case-study level. Existing research does not address how established companies can use their data assets to develop new data-based business models.

Thus, in this paper, we focus on better understanding possible business models that companies can use to create and capture value from their data assets. To do so, we empirically study the various ways in which firms monetize data they own or have access rights to. We focus on the key choices that companies make when setting up data-based business models (i.e., business transactions' content, structure, and governance) (Osterwalder, Pigneur, & Tucci, 2005; Zott, Amit, & Massa, 2011). We focus especially on the process that companies use to craft value-creating market offerings (Saxe & Weitz, 1982; Terho, Haas, Eggert, & Ulaga, 2012) or how they monetize data in the marketplace. Our interest lies in the reasoning companies apply when choosing from different business-model and sales-process options and in how they have, or intend to, design their operations both to create and appropriate maximum value for their customers and themselves. We also focus on identifying the challenges and constraints that prevent data monetization and shape data-based business models. To our knowledge, our study represents the first to shed light specifically on the activities that drive data monetization in established middle-sized and large firms.

Accordingly, we conducted qualitative in-depth interviews with 37 individuals from 24 large business-to-business (B2B) and business-to-consumer (B2C) companies representing 11 industries. We selected companies that we knew to have engaged in, or planned to engage in, data monetization. The interviewees had much knowledge in data utilization and commercial data use and had titles such as chief digital officer, chief information officer, digital transformation manager, director of new ventures, and sales director. We analyzed the interview data thematically to find recurring issues and possible best practices and to categorize them in accordance with our study's theoretical foundations.

We expect both established companies planning to engage in data monetization ("What options do we have?", "What routes should we take?") and individuals in society who care about data privacy to find our research questions relevant. A growing pool of private citizens and various advocacy groups share concerns about how companies use people's private data (e.g., Crain, 2018; Sharma & Gupta, 2018; Solon, 2018b), and we urgently need to understand the companies' standpoint in this discussion. We believe that understanding companies' perspective will narrow the gap between the public and companies and enrich the discussion by bringing in corporate views on data monetization. Many companies behave secretively in this respect even if they act legally and in ways that align with the general public's moral views on the matter. On the academic level, our study contributes to various research areas, such as big data-focused studies in information systems, studies on business models and multi-sided platforms, and

studies on data privacy and security. With this paper, we focus on making it easier for companies, researchers, governments, and other stakeholders to engage in and understand ethical data monetization.

2   Theoretical Background

2.1   Data as an Asset

While many use the words data, information, and knowledge interchangeably, they represent distinct concepts. Ackoff (1989) describes data as a set of symbols that represent event and object properties, whereas Liew (2007) refers to data as unrefined and unfiltered information or a set of discrete objective facts about events. Alavi and Leidner (2001) define data as facts or raw numbers and information as processed and interpreted data. On the other hand, they define knowledge as actionable information or meaningful links between the information and its applications.

Data is a non-rivalrous experience good (Quah, 2003; Rayna, 2008; Pantelis & Leiponen, 2013). Non-rivalrous means that everyone can access the good (in this case, data) when revealed. Furthermore, non-rivalrous means that multiple actors can exploit one single data set at the same time and that using it does not make its value disappear. Data's value depends on the context, situation, and time. While initially collecting data—an intangible information good—can involve great cost, copying and disseminating it typically do not. Experience refers to the good's use value. Thus, with data, one can realize its value only when one uses it; otherwise, one cannot define its value. Furthermore, when one reveals data, its value reduces rapidly. Data's value as such varies by the way one uses is, how one combines it, and whether one can provide it at the right time and place.

From an organizational perspective, data may come from internal or external sources (Negash, 2004). Organizations acquire internally sourced data from their existing IT systems that concern either natural persons (e.g., CRM data, service-user data) or business operations (e.g., ERP data, sensor data) (Haddare & Elragal, 2015). In contrast, organizations acquire externally sourced data, which we can divide further into bought and open data, from outside their borders (Hartmann et al., 2016). Open data is downloadable, machine readable, and structured without prior processing: prime examples include open social media data and data that public actors make available (Lakomaa & Kallberg, 2013).

Various writers refer to data as raw material for a new way to create value and identify situations in which it may be valuable. Examples include optimizing transportation services, personalizing the way in which users experience digital services, and improving service solutions for heavy machinery (Parmar et al., 2014; Rossman, 2016; Spijker, 2014). Improved analytical capabilities and cloud computing also drive the change: expanding capacity in terms of storing, analyzing, and sharing data have fostered innovation in data-based services (Najjar & Kettinger, 2013; Parmar et al., 2014). The myriad ways in which organizations can exploit data often emerge unexpectedly, and the data's value depends on whether one aggregates or anonymizes it. As an example, the various smart city and smart traffic scenarios rely on data ecosystems, which themselves often rely on open data (Lindman, Kinnari, & Rossi, 2016), that allow for city planning, route planning, road tolls, insurance optimization, or citizen monitoring (e.g., see Rathore, Ahmad, Paul, & Rho, 2016).

2.2   Approaches to Monetizing Data

Data monetization refers to capturing data's value (Teece & Linden, 2017). Najjar and Kettinger (2013) suggest that companies can monetarily benefit from data by selling it, trading it, or optimizing their operations and, thereby, reducing their costs. Companies largely have two monetization options: 1) improving existing processes and products (internal uses of data) and 2) creating new offerings (external uses) (Wixom & Ross, 2017). Significant documentation on how companies can use data internally exists; thus, academics and practitioners alike have tended to focus on its internal use (Najjar & Kettinger, 2013; Thomas & Leiponen, 2016; Woerner & Wixom, 2015). As such, data monetization addresses data's unique propersties as an intangible asset that holds opportunities for creating and capturing value through novel ways to exploit gathered and organized data (e.g., Constatides, Henfridsson, and Parker (2018) discuss how John Deere used data that it gathered from farmers' tractors).

A business model is a conceptual tool that expresses a specific firm's business logic (Osterwalder, Pigneur, & Tucci, 2005). It encompasses the content, structure, and governance of transactions that purport to add value (Zott et al., 2011). As Osterwalder and Pigneur (2010) argue, although a business model reflects various business-related decisions, the value proposition represents its key element. The

value proposition covers decisions regarding the cost structure, the main production elements, revenue streams, and customer segments. Researchers have previously explored some of these aspects in relation to data monetization (Najjar & Kettinger, 2013; Thomas & Leiponen, 2016), but we lack research data that demonstrates insufficient business models or the most effective business models for selling data itself. Furthermore, we lack agreement on what to call them: big-data, data-based, or data-driven business models (Schüritz, & Satzger, 2016). From the literature, we identified three types of business models for creating and capturing value from data that differ in the extent to which they increase its value: selling data, selling analyses, and selling data-based services. These types differ based on whether the buying company will obtain ownership of the original data or not.

2.2.1   Selling Data

Arguably the most straightforward way to monetize a company's data externally involves selling it directly to another party and, thus, granting its ownership to the buyer (Spijker, 2014; Thomas & Leiponen, 2016). Organizations provide raw or aggregated data for others as such, but parsing, cleaning, or cataloging it could increase its value (Thomas & Leiponen, 2016). Other potential raw data buyers include data aggregators who combine and analyze data from multiple sources: this business model has improved in viability as data masses and sources have multiplied (Brown, Chui, & Manyika, 2011; Thomas & Leiponen, 2016).

Examples of companies selling data include:

- Vodafone sells anonymized network data to navigator producer TomTom. Vodafone has real-time, location-based data about its customers that TomTom buys to better understand how people and vehicles move and, thus, optimize navigation suggestions. (Parmar et al., 2014)

- PatientsLikeMe sells anonymized data from its social media platform to pharmaceutical companies and companies that manufacture medical devices. The data concerns the platform users' experiences related to illnesses and treatments, which helps the partners to develop more effective and better-suited solutions. (PatientsLikeMe, 2016)

- Bundlers sell real-time location data from phones to marketers in the US (Valentino-DeVries, 2018)

2.2.2   Selling Analyses

Another approach involves selling data-based analyses but restricting access to the original data (Spijker, 2014; Thomas & Leiponen, 2016). Companies that purchase the analyses (insights or benchmarks) generally use them on a one-off or continuous basis to better understand consumers, competitors, production elements, or business environments and to improve decision making. As a benefit compared to selling raw data, analyses better protect data's privacy and security as the data's owner does not change (Spijker, 2014). On the other hand, more complex or specific the analyses typically have fewer potential customers. Moreover, buyers can use the raw data in more versatile ways by combining it with other data sources.

Examples of companies selling data analyses include:

- The pharmaceutical distributor Tamro provides drug manufacturers and other suppliers with insights into consumers' spending on their own products in specific locations. Tamro also offers pharmacies information about their sales compared to their competitors' sales (Tamro, 2017a, 2017b).

- Barclays provides a service for SME companies that provides them with information about their revenue inflows and outflows and analyzed data about their payments and transactions. Customer can also compare their own data with that data from similar businesses in similar locations (Barclays, 2016; Spijker, 2014).

2.2.3   Selling Data-based Services

Some companies create new services to scale how they deliver data using multi-sided business models, which they usually implement via dashboards or similar digital interfaces (Najjar & Kettinger, 2013; Spijker, 2014; Whitmore, 2016). The data comes from their services' users, and companies monetize the data. Some users pay to use such services by providing data (Spijker, 2014). In such a business model, customers do not own the raw data and can see its nature only to a limited extent.

Examples of companies selling data-based services include:

- Google's smart thermostat Nest provides data on users' energy consumption to electricity utilities. Nest does not share the user's raw data with the utilities, but it does enable them to balance the user's energy grid (Dillet, 2014).
- Facebook sells advertising space by enabling customers to target their advertisements to specific user groups based on their user data (Matsakis, 2018).

2.3 Contingences and Constraints on Developing Data-based Business Models

Despite the various opportunities in data monetization, various contingences and constraints affect efforts to develop data-based business models. The existing literature stresses organization type, position in the value network, data characteristics, data privacy and data security in particular.

2.3.1 Organization Type

Organizational stability and consistency in many established organizations constrains data monetization, which usually requires innovation, proper organizational alignment, and a good understanding of data. Hence, many organizations need to put new processes and resources in place—particularly if they expect the new business to grow significantly compared to the existing business (Bean, 2017; Spijker, 2014). The discussion also concerns suitable organizational types. Spijker (2014) argues that new data-related projects need to have strategic importance because they require organizational transformation. Innovation theories in large corporations support this notion: new business ventures require a dedicated team, autonomy, informal networks, room for experimentation, and risk taking (McDermott & O'Connor, 2002; Stringer, 2000).

2.3.2 Position in the Value Network

A company's position in its value network, as a contingency, affects its ability to start monetizing data; in most cases, moving along the value chain or the value network represents the most natural course of action (Spijker, 2014; Thomas & Leiponen, 2016). This movement resembles forward or backward integration: combining different data sources from the same value chain or network creates added value (Amit & Zott, 2012). Thus, organizations with access to various or vast data sources can arguably monetize data the most successfully (Thomas & Leiponen, 2016). Nevertheless, a company can still monetize data even if the paying customer is not part of the selling company's value network as long as they share another common factor, such as a shared final customer or a shared production method (Spijker, 2014).

2.3.3 Data Characteristics

Data quality often represents an initial issue that companies face when starting to create a new data-based offering (Spijker, 2014; Wixom & Ross, 2017). Other key quality factors include accessibility, accuracy, timeliness, completeness, consistency, relevance, and liquidity (Yousif, 2015). Quality problems tend to arise if a company creates and stores data in silos without a proper data infrastructure (Spijker, 2014). Companies need sufficient data quality to be able to combine different data sources, create access to the data, and to increase the customer's trust in the data (Wixom & Ross, 2017).

2.3.4 Data Privacy

Companies need to consider privacy concerns when data includes personal information about customers or personnel. Many companies do not monetize data due to reputational risks (Thomas & Leiponen, 2016). Indeed, data leaks, which the Cambridge Analytica scandal exemplifies (Solon, 2018a), receive significant attention from the public, and lawmakers scrutinize personal data that companies use commercially more intensely. Furthermore, many studies have shown that private users rarely internalize privacy agreements (Debatin, Lovejoy, Horn, & Hughes, 2009; Jensen, Potts, & Jensen, 2005; Sheehan, 2002), which increases the risk that companies will lose customers' trust if they share data with external partners. Laws and regulations on data privacy impose limits on what data companies can collect and exploit but also clarify the rules for all players.

2.3.5   Data Security

Data security relates to confidentiality and the ability to protect data from theft, errors, and accidental destruction. A company may need to secure data from access by external parties to protect confidential business information. As such, authorization, access, auditing, and encryption become relevant factors when a company handles business-sensitive data (Yousif, 2015). Matters about trust and ownership may foster conflicts between sellers and buyers in cases of data monetization (Thomas & Leiponen, 2016). Therefore, data security imposes significant constraints and risks on the type of business model that a company can use in the monetizing process.

3   Methodology

3.1   Research Approach and Data Analysis

Research on data monetization and data-based business models has primarily appeared as conceptual papers and commentaries (Opresnik & Taisch, 2015; Woerner & Wixom, 2015), bibliometric and review studies (Chen et al., 2012), and single-case studies (Najjar & Kettinger, 2013). As such, to open up an empirically more comprehensive perspective on this new phenomenon, we adopted a qualitative research approach. Given their exploratory nature, qualitative methods allow one to obtain deeper insights into weakly understood issues or shed new light on well-known phenomena (Birkinshaw, Brannen, & Tung, 2011; Ritchie & Lewis, 2014). In our study, we combine exploratory and descriptive approaches to provide a wide-ranging view of data-monetization practices in large and middle-sized companies.

Given that we lack empirical studies on data monetization strides that middle-size and large firms have made, we decided on an inductive research approach since it suits exploratory research, particularly when one does not sufficiently empirically understand a nascent phenomenon in a real-world setting (Golafshani, 2003; Ritchie & Lewis, 2014). We did rely on a semi-structured interview guide to create a mid-range theory on data monetization that accounts for contextual variations and local systems (Reeves, Albert, Kuper, & Hodges, 2008). In inductive research, one begins by collecting data and analyzing the results to uncover emerging themes (Saunders, Lewis, & Thornhill, 2009). Thus, evidence serves as the genesis for any conclusions (Ritchie & Lewis, 2014).

Data collection and analysis often go hand in hand in inductive research (Gioia, Corley, & Hamilton, 2013): the analysis starts as soon as the first research evidence emerges (interviews and workshop discussions in our case). We documented the interviews by audio-recording them if we had interviewees' permission; otherwise, we took notes during them. After analyzing the notes or transcribing the recording for each interview, we categorized the data into emerging themes. Researchers commonly perform such an iterative analysis when conducting exploratory research (Saunders et al., 2009). We used open and selective coding to analyze the data when we summarized and categorized the results. We used similar methods to those associated with grounded theory, including memoing, constant comparison, and selective coding. A recent meta study on grounded theory use in information systems (Wiesche, Jurisch, Yetton, & Krcmar, 2017) concluded that such studies contribute to the literature "by developing pre-theoretical models and rich descriptions of new phenomena" (Wiesche et al., 2017, p. 697).

3.2   Data Collection

We wanted to investigate various businesses to comprehensively overview data-monetization activities regardless of industry. We also wanted to include companies in different data-monetization phases to enrich our knowledge about the barriers involved in monetizing data and about successful and non-successful practices and decisions. Thus, we sampled large and middle-sized companies engaged in both B2C and B2B business; companies that possessed data on service users or consumers; and companies with data on, for example, buildings, processes, or machinery. Since we focused on studying how companies monetize data collected from existing business operations in this paper, we lacked the scope to consider small-sized companies and startups (for a review of data-based startup businesses, see Hartmann et al., 2016).

With help from two IT companies with substantial experience in data-monetization projects and good access to middle-size and large firms interested or engaged in data monetization, we contacted 44 companies with ongoing or recently ended data-monetization projects. Of that number, 24 companies representing 11 different industries agreed to participate in the study. Given we focused on exploring a

nascent phenomenon, we collected two types of case evidence. First, we interviewed representatives from the case companies; we conducted 37 interviews in total. We targeted informants with knowledge about their companies' data and IT management, new business ventures and development endeavors, or sales or general business management. We provide the interview guide in Appendix A. Second, we organized three discussion workshops among selected companies from the sample, which involved 17 company representatives in total. Accordingly, we built a comprehensive picture of data monetization in the focal organizations and industry specificities in data monetization. We collected all data in 2017. We provide details about the case companies in Table 1.

Table 1. The Sample Companies

Company	Industry	Number of interviewees
A	Construction	4
B	Construction	1
C	Energy	3
D	Energy	1
E	Finance	1
F	Finance	1
G	Finance	1
H	Forest	1
I	Healthcare	1
J	Healthcare	1
K	ICT	5
L	ICT	1
M	Industrial services	3
N	Industrial services	1
O	Manufacturing	1
P	Manufacturing	1
Q	Manufacturing	1
R	Media	1
S	Media	1
T	Media	1
U	Media	1
V	Retail	1
W	Transportation	3
X	Transportation	1
Totals		
24	**11**	**37**

All the companies had significant operations in Finland and most also internationally; therefore, they operated under E.U. law. The General Data Protection Regulation (GDPR) (approved in 2016 and implemented in 2018) sets several requirements for and limitations on data monetization in E.U. countries. Lawmakers implemented the GDPR to harmonize data-privacy laws across E.U. countries and to increase E.U. citizens' control over their personal data (European Commission, 2018). The GDPR gives citizens (data subjects) the right to be forgotten, to request organizations to erase their personal data, and to retrieve their own personal data and share it with another service provider. It also gives companies clearer and more consistent rules than they had before and also imposes harsher penalties on organizations that do not comply with the laws and neglecting data privacy. Indeed, sanctions on companies that do not comply with the regulations can reach up to four percent of their annual revenue or 20 million euros (European Commission, 2018). All the companies in our study knew about the upcoming legislative change and had begun making adjustments to their data-management operations accordingly.

The Payment Services Directive 2 (PSD2) affected the banks and financial institutions in our sample. The directive purports to protect consumers making online payments, promote open banking, and make cross-border payment services safer (European Commission, 2015). In terms of privacy, PSD2 includes a new regulation about handling and protecting consumers' financial data.

The interview guide included questions in the following categories: 1) to what extent the interviewee's company and industry used data, 2) whether the company buys data from other organizations (if relevant, then what kind of data, from whom, and on what basis), 3) whether the company sold data to other organizations (if relevant, then what kind of data, to whom, on what basis, and how), 4) new data-based services and platforms (if relevant, then what kinds of services, to whom, and on what basis), and 5) other relevant issues. We conducted the analysis by identifying words, sentences, and themes through text analysis. We conducted the analysis in two stages so that, after collecting and analyzing data in the first round, we collected more data that we then analyzed together with the previously collected data. The first and second authors conducted the coding process, but we all discussed the results several times to reach a consensus.

4   Results

In general, all the companies used data to enhance internal processes and decision making and to improve their current products, services, marketing, and other things. However, they differed in the extent to which they used data: some companies were already strongly data driven and had intelligent and contemporary IT systems, whereas others still operated in legacy environments with siloed databases. Over one third of the studied companies already had data-based offerings, though most were not comprehensive and represented nascent businesses or small-scale experimentation. A few companies had thriving, revenue-generating data-based offerings with sales targets and dedicated sales personnel. The remaining companies had voiced interest in monetizing data but did not have data offerings up and running. We also found interviewees from these companies in particular who criticized data monetization.

Those companies that had more advanced IT systems with high-quality data (media companies in particular but also among advanced manufacturing firms with strong service businesses) used the data for a wider number of internal tasks and processes. Interviewees for the media companies mentioned that data directly or indirectly enabled nearly half their annual revenue. Furthermore, the advanced manufacturing firms had gradually enhanced traditional products with digital features. Through that, they had created new features to their existing businesses, which also had started to generate an increasing amount of data. However, due to this increase, the firms faced pressure to develop their data capabilities (i.e., through developing master data management, providing application programming interfaces (APIs), and forming a dedicated data science team). Over time, the increasing data capabilities allowed the companies to further start integrating external data and to consider new ways to use their internal data assets. Better IT infrastructure enabled companies to seek out for alternative ways to benefit from their data: the companies struggling with their IT infrastructures (i.e., companies that had constraints with sharing data internally and constraints in information flow both in and between business function) mainly focused on obtaining better IT capabilities and using data internally, while companies with more advanced data and analytical capabilities had started to exploit new data-based business models and business opportunities. We also found that properties of the data in legacy systems (i.e., the age, maturity, and sheer number of old systems) and acquisitions had slowed down the development of data utilization in several companies. Thus, the nature and maturity of the IT infrastructure in companies affected the opportunities they had to monetize data and the challenges they faced. In Section 4.1, we present the results concerning the companies' approach to data monetization and the developed data-based business models.

4.1   Data Types that Companies can Monetize

Most prominently, among the companies we examined, we found that access to certain data types offered various categorical benefits in terms of monetizing data. A company that had one or several digital elements in its business model was more ready to create data-based offerings. Rather intuitively, a company seemed to benefit if it had digital products or services, customer relationships, and sales channels.

First, production systems that automatically collect digital usage data (e.g., E.U.-mandated smart energy meters and bank transaction systems) benefitted energy and finance companies and hindered

construction and healthcare companies (e.g., legacy systems with non-standard data). A director in a construction company pointed out that companies involved in building urban environments are the least likely to participate in data-centric IT business, such as access control. Another construction-company representative pointed out that, unlike shopping-center operators, for example, his company did not sell data-intensive information-systems services:

> *The problem is that we are not there when they [buildings] are used. Our competence is putting together a large construction project. Oftentimes these ideas [about data monetization] deal with developing services for living in the buildings, but we are not there.*

Interviewees from the construction companies also pointed out that no usage data currently accumulates automatically and that monetizing data would need a separate sensor strategy, which decreases data-business initiatives' attractiveness. Some ongoing data-led business initiatives among the companies did accumulate data automatically, such as through a builder's safety equipment-monitoring system. In the healthcare context, informants complained that, although patient medical records contained a lot of information, organizations commonly do not automatically generate data about complying with prescribed medication, rehabilitation, and recommendations concerning eating and exercise.

Second, customer data is valuable, and companies with digitalized customer relationships enjoy the benefits from storing and organizing it in a useful way. An interviewee from an ICT company operating in both consumer and corporate telecommunications reported stronger data-progress business progress in the organization's consumer business unit. The interviewee acknowledged that the fact that the consumer business unit had organized its data vault better had helped it proceed faster in monetizing data even though the company had richer data on corporate users, wider corporate uses cases, and clearer data-monetization business cases. For example, the interview mentioned consumer location data, which one needs to clean a lot, as data that the organization had organized particularly well and that had led to successful monetization initiatives. We can see a similar situation with raw open government data: while the raw data has little or no value, aggregators can combine and clean it and make it valuable for service and app providers (Lindman et al., 2016).

Third, digital sales and delivery channels yield accurate data that companies can monetize in new ways. A retail-industry informant reported that his company had comprehensive and detailed real-time point-of-sales data from its transaction systems and that, unlike its competitors, it had insourced running of the point-of-sales systems and online store checkouts and, thus, had all the relevant click stream and online store data. The company had launched a value-added service in which a customer who bought a product from selected product categories could receive immediate digital follow-up offers directly related to the purchase, such as tools, installation services, transport, spare parts, and training.

A respondent from an ICT company pointed out that selling smartphones and data subscriptions allowed it to access the user interface that customers used to manage their data services and add data prices to the subscription-based invoicing arrangement it had with them. Having this relationship enabled the company to add other data-based services (e.g., to monitor real estate, control energy efficiency, and view family members' data usage) onto customers' "phone bill".

4.2   Inducers and Triggers for Data Monetization

All the sampled companies had interest in monetizing data, but their offerings varied in maturity. For the majority, we found that external pressure had instigated their data-monetization initiatives. In particular, respondents most often mentioned abstract pressure from the market and competitors and the threat of disruption, but some also mentioned competitive threats in the market. Several quotations exemplify this latter sentiment:

> *We cannot afford NOT to make these data-led experiments.*

> *We will fall behind in the competition if we don't build new services based on data.*

> *If we do not start utilizing our data better, this company will not be here in 10 years.*

We less commonly found data-monetization projects that originated from a clear need that customers had expressed. As a notable example, a director from a company that provided industrial services described how the company created its data-based offering as a direct response to customer requests. According to the interviewee, since the company had identified a clear customer need, it could easily begin developing the service.

The changing legislation encouraged many companies in our sample to consider monetizing data. Most knew about the GDPR and PSD2 and complied with the requirements, and many considered the new directives a business opportunity due to the clearer rules about using data. For example, two interviewees said:

> *Currently no other company uses our data, but PSD2 will change that. Direct monetization has not thus far been possible, but PSD2 will encourage data sharing and even monetization.*

> *The regulatory challenges have never been a real issue. Like we would have had to raise our hands and say we simply cannot do this. You'll always find some way to deal with them.*

Companies in other industries such as manufacturing, construction, and electricity lacked clear market pressure and incentives to more broadly use data. Interviewees for these companies said that they mainly focused on internal data projects (i.e., making their current business more valuable with data). Some companies, particularly companies who held sensitive and unique data (e.g., in healthcare), had evaluated the risks from monetizing data as outweighing the potential benefits. Specific risks included losing customers' or service users' trust, harming their reputation, and damaging their brand image. Interviewees from the forestry and logistics firms also said that decreased revenues had pressured the firms to increasingly use data to survive the changing market.

4.3 Control of Customer and Partner Selection

Some companies (mainly the media organizations) in our sample sold data as is or in an aggregated form. Their offerings took various forms. Some companies provided aggregated data as a one-off file transfer (the simplest form to provide data), whereas some offered more continuous data stream with subscription services and priced APIs that provided live updates and access control. One company had a pilot project with a research company to which it sold anonymized but detailed data on consumers' online activity. Most companies, however, consciously avoided selling raw data to protect their valuable assets and reduce legal risks.

Notably, the companies that sold raw-data feeds offered them only privately to selected partners. In this way, they reduced (though did not eliminate) the chance that someone would inappropriately use the data. For example, we found this approach in a healthcare company that did not share its data directly but provided pharmaceutical companies with anonymized and aggregated data on who used their products. Furthermore, a manufacturing company carefully selected customers that it sold its data to but did so either as a raw-data feed, discrete dataset, or a ready-made analysis depending on the customer's use case. A representative for a media company referred to its strategic partnership with another company from which it obtained anonymized data that it integrated with its own data:

> *We don't intend to share our data with just anyone, and we'd rather choose a couple of partners with care. They should have a clear use case for the data and enough analytical power and skills.*

Even though interviewees indicated that their companies did not see data security or data privacy as issues in cases of data-providing services or platforms, their companies still wanted to carefully choose the parties they allowed to their platforms—even if their platforms benefitted from a wide variety of partners and collaborators. In particular, some expressed concerns in terms of what partners would do with their company's data and their alignment with its brand. We note that interviewees mentioned these concerns when explaining why their firms lacked a data-selling solution. For example, some interviewees said:

> *Pretty quickly the brand gets into the picture. If someone were to sell a shitty service in our name.*

> *My view is that, especially at the beginning, the platform cannot be completely open due to the sensitivity of the data. And because of that, we need to control the quality of the services built on top of our platform. We wouldn't want to open it to the whole world and lose our reputation if something didn't work properly.*

In some cases, a company chose a partner not due to the latter's superior data or related skills but due to the non-data related history between the two. For example, a manager in an energy company said:

They [existing partners] don't have anything to do with data. They just do up-selling. Maybe we'll notice that another partner would be better [for developing data-based services with] than those whose services we're currently selling.

4.4   Prioritizing Existing Business and Customer Relationships

The majority of the studied companies did not monetize data as a common business practice. Most had had pioneering data offerings. The data business did not constitute a major business at the time we conducted the interviews. Therefore, most informants did not see much promise in data. For example, a team leader in a transportation company voiced the concern: "Well I guess it could go so that we created a new firm for these data projects. But why wouldn't I just focus more on our own [existing] business?".

As many informants noted, the opportunity costs related to data initiatives delayed experimentation with new data offerings. Moreover, new initiatives regarding selling data tended to lack priority. The potential benefits of data monetization and the effort it required were more likely to accrue in internal processes and projects. For example, two informants said:

Currently the potential business in data utilization lies more in internal processes and existing products than in data monetization.

For us, it is more important to create additional value for our customers through new services, than just to get a tiny increase in revenue by selling data [to somebody else].

Many informants made such comments due to their concern about what value they added for their current customers. Indeed, they emphasized that most revenue comes from a firm's core business and not side projects. Therefore, they focused primarily on internal issues and creating added value for existing customers. Some said that they even provided data to their existing customers but without any fee or cost—they provided the data as an additional service element or to keep a customer happy.

Another related challenge concerned the potentially difficult relationship between potential new data business and a firm's existing business. In some cases, the evident new data business challenged or even competed with a firm's existing business. However, most companies did not experience this issue, and we found that, in several cases, the data offerings targeted a completely new customer base. For example, two informants said:

The business logic would need to change because otherwise [with data offerings] we would expose our secrets to our existing partners. We cannot corrupt our existing business.

If we want to do some of those things [data business opportunities], we cannibalize something that already exists. And it is a question of the right timing. Because someone will do that. Or do we just accept the fact that some businesses will just fade away?

4.5   Frustration with the Lack of Reference Cases and Prices

According to our results, the majority of companies struggled with how to value their data. In general, they attached a higher value to consumer data than to other kinds of data not least due to its more evident and readily available potential use cases. For example, compared with the other case companies, manufacturing companies clearly perceived their data as lower in value because it lacked clear direct applicability.

The companies also perceived raw data as having rather low value but that refining it or combining it with other data sources might give it added value. It seems that many companies in different industries increase value by refining data. Furthermore, some interviewees pointed out that not data but the resulting algorithms and analyses created the real value. On the other hand, they found raw data more valuable when it was unique to the business environment: the value increased if no other actors possessed and provided similar data. Thus, we can see that companies valued data along a continuum: some companies viewed data as their most valuable asset, whereas others had difficulties in seeing any value in it, which the following quotes show:

I don't think that our data is valuable to any other organization.

And the data that is not that relevant for us…. It's so boring, would anyone want to buy it?

Naturally, given that the companies found it hard to define data's value, many also struggled to price their value-based offerings. The biggest difficulty concerned the lack of reference prices in the market. A director in an ICT company described the problem as follows:

> *We have four pilot customers that would like to test it, but we have not come to an agreement on any kind of pricing model. – The problem is the different views on the value of the data. Because pricing usually depends on information about other similar products on the market. Now, there are no similar products on the market.*

Relatedly, we found that companies found it hard to define the possible market's size. A director in a media company emphasized the need for courage in saying:

> *There are different business models with monthly subscriptions, SaaS and research-based models, for example. We have pondered on the models and pricing of our products for a few years now, but I cannot even describe what the market is, how big it is or who the players are. But we don't let this hinder us too much. We're fulfilling our customers' needs and they are paying for our data. So, I think you just need courage to go forward, and the market will hopefully become clearer as you proceed.*

Many interviewees mentioned a lack of demand and customers' lack of openness to new solutions. Companies with existing data offerings had addressed the issue of creating demand: customers have to believe that the data offering provides value and that they need it. One third of the companies pointed out that it does not suffice that the data solutions create added value: customers also have to be ready to use them. Customer maturity emerged as an issue in almost all the represented industries.

On the other hand, some interviewees admitted that, although they their companies had an interest in developing data offerings, they did not systematically map the potential customers and talked with them serendipitously. A team leader in an ICT company said:

> *There is demand and we get requests from customers, but our salespeople have to say that we cannot offer that at this moment. Maybe the reason why we still don't lose customers is that none of our competitors are offering anything like it, either.*

Many informants raised questions about monetary returns and data monetization's profitability. Moreover, given the lack of clarity about potential demand and pricing, companies had scarce business-case calculations. Some informants pointed out that monetization would never provide sufficient revenue for established companies and others that their companies had no clear financial incentive to create data-based offerings in the first place. In contrast, informants from industries that had already undergone digital transformation, such as banks, financial institutions, and insurance companies, indicated that they had already moved to monetize their data-based knowledge about risk through various digital insurance offerings based on, for example, data collected from Internet of things (IoT) devices (Baesens, Bapna, Marsden, Vanthienen, & Zhao, 2016).

4.6 Organizational Enablers and Barriers for Starting Data Monetization

Some interviewees thought that their organization's organizational culture prevented data monetization by discouraging new business development and risk taking. Many interviewees whose company lacked such solutions pointed out that their company's inability to identify and act on business opportunities in particular hindered monetization. The interviewees indicated that they could already see these problems in how their organization used data internally. For example, one interviewee said: "If there is no culture or encouragement to share information even between projects, it is hard to monetize or even to utilize data company-wide".

We found that all the industries we examined had a need for management focus and support and that traditional business-case calculations do not help data-monetization projects. Therefore, as the monetary benefits can appear vague at first, interviewees perceived stronger support from management as important. For example, two interviewees said:

> *This is the kind of thing that doesn't go bottom-up. Won't work. It needs to be top-down, which requires vision from the leaders to see where the world is going.*

> *I'm frustrated sometimes. I've noticed that my and the top management's thinking can be so different. You need to be visionary to believe in the bigger picture that seems to be evolving.*

The interviewees believed that several measures indicated proper management support, such as investing in venture accelerators or programs, setting up completely new business ventures, and forming independent teams that focus on data offerings. An interviewee from an energy company said:

> We've had hackathons and things, but I have not heard of anything concrete coming out of them. They are one-timers. They're lacking continuity. If you believe in something, you should invest in it and continue.

> This [data monetization] requires a completely independent team, and it wouldn't necessarily even be part of us. It would be a company of its own.

Furthermore, many interviewees believed that top management should decide what exactly the sold service should be like. A director in an ICT company described the need for a top management vision and engagement as follows:

> Maybe we should be more active and say, we can do this for you guys [customers]. We do it so and so and it costs so much. I mean it should be productized. It would require us to decide that we are going to build this kind of data-led business that is not about process optimization. But we would optimize someone else's business. You should build a team that would take care of it. You first need a vision: what, how and for whom.

Relatedly, many informants pointed out the change in attitudes to IT tools and resources: they no longer constituted a barrier or bottleneck to using data but rather enabled it. Although their organizations still had demand for technological superiority and skillful data scientists and analysts, the informants stressed the need for skills to create market offerings from data. These skills related particularly to the ability to understand and communicate data's value from the customer's perspective, to innovate new data-based offerings that serve a customer need, and to operate in the interface between business and IT. Informants from different companies said:

> The development of the technology related to data is certainly a vital necessity, but it will come, and it's not rocket science at the end of the day. Maybe the biggest challenge is that the business guys should know more about data.

> A bigger challenge than regulation is the ability to innovate. Or service-design innovation is the right word. Technical innovation we can do.

> It's more about the fact that there are limited resources both in IT and in business. It doesn't help if you take twenty fellas from outside if they don't understand our business. That's a bigger dilemma [than technical capabilities].

4.7   Regulation and Data Properties are Shaping Data Monetization

The interviewees referred to regulation as an evident factor that restricts data monetization in a majority of the companies. In particular, they noted it prevents companies from directly selling customer data. Furthermore, the law prevented some companies from engaging in any other business than the one in which they operated, so they could not feasibly sell data as a new business venture. The changing regulatory and legislative environment and the GDPR made them more careful, especially in companies dealing with consumers. It seems that, although a European legislative effort, the GDPR will have global regulative effects. For example, an informant said:

> GDPR slows down investments related to data. The lack of precedents increases caution since no one wants to be the first one to test the boundaries.

Many companies, particularly the ones that held sensitive and unique data (e.g., finance and healthcare companies) had evaluated the risks from monetizing data as outweighing the potential benefits. Specific risks included losing customers' or service users' trust, harming an organization's reputation, and damaging its brand image. For example, two informants said:

> The position of trust we have with our customers is so strong that potential monetization deals are not even considered, since their business value could not be worth the risks in terms of decreased trust.

> Data is a matter of honor for us, so we do not want to test the boundaries of our customers.

> *We are a little anxious about what that means from the perspective of GDPR even though we know that if we anonymize and aggregate the data, there is no problem. But the brand risk is something we evaluate. If you logged on [to a data broker's service] and saw our logo there… that's not something we'd want.*

5 Results Summary: Data-based Business Models and Paths to Data Monetization

Data monetization, although an unknown opportunity for many traditional industries, could offer companies new ways to create completely new service offerings. According to our findings, many companies start monetizing data by offering new services to existing customers in part because they tend first to experiment with single customers and partners. Many interviewees believed that this approach allowed companies to more easily evaluate whether or not they could add value from monetizing data and to better control data. However, the most advanced companies focused on a scalable approach and had created data-based services that they offered to completely new customers. These approaches seemed to provide more sustainable and financially attractive businesses for the selling companies.

Thus, we can categorize data-monetization business models based on two dimensions: the type of customers the offering focuses on and the offering type. We can categorize customers into three types: current customers, actors in the current value chain, and anyone interested to buy. The offering type follows the categorizations found in the previous literature (Spijker, 2014; Thomas & Leiponen, 2016), which we refine further in this study. Companies can directly sell data, analyses, or data-based services. In both dimensions, the latter categories typically possess more opportunities for scaling the business due to the number of potential customers and effort required to make a sale. These categories create a 3x3 matrix and nine specific business model types that vary in requirements with respect to data properties, data infrastructure, data governance, data management, and data capabilities. We summarize these different business models and identify the case companies that had applied or considered applying such business model in Table 2. The table distinguishes between business models per customer type and shows how the initiatives in the companies we studied skewed towards current customers and value chain partners.

Table 2. Business Model Variations for Data Monetization

Customer type/ offering type	Current customers	Actors in current value chain	Anyone
Selling data	Selling or providing data* to current customers Companies: C, O, R, S, T, U, X	Selling or providing data* to current suppliers and partners Companies: C, E, D, G, H, I, J, L, N, O, R, S, X.	Selling data* to data brokers or data integrators
Selling analyses	Selling analyses on the business environment to current customers Companies: O, R, S, T, U, X	Selling analyses on the business environment to current suppliers and partners Companies: C, J, O, R, X	Selling analyses on the business environment to market research agencies Companies: O, T
Selling data-based services	Selling a service that provides current customers relevant signals on the business environment Companies: K, M, O, R, S, T, U, X	Selling a service that provides current suppliers and partners with relevant signals on the business environment Companies: C, E, K, M, O, Q, R, T, U, X, W	Selling a service that provides new customers relevant signals on the business environment (B2B or B2C) Companies: O, T
Notes: * Non personal, anonymized, and/or aggregated data.			

According to Piccini, Hanelt, Gregory, and Kolbe (2015), the traditions attached to managing physical business models make companies careful when selling data to their current customers and other stakeholders in their current value chains. The way that traditional business models treat competition and markets on the one hand and privacy, security, and legal liability on the other makes companies wary of openly selling data to anyone. However, many companies we examined used data as an additional service feature that they provided to current customers or partners either for free or for a cost. Regardless

of the industry, we clearly seemed to have found certain roles in the value chain that had a better position for monetizing data than others and that data monetization might be becoming a vertical positioning exercise. With regards to Piccini et al.'s (2015) findings, our results also indicate that the further a company develops an offering from raw data, the more likely the company will aspire to sell it openly to anyone. Considering that data brokers aspire to sell others' data directly, our results suggest that either data owners increasingly do not sell to brokers or that data owners clearly need data brokers since the former shy away from selling raw data to anyone.

As Zott et al. (2011) and Bharadwaj, El Sawy, Pavlou, and Venkatraman (2013) point out, business models that emerge from digital strategies face the limitations of traditional business models. From our interviews, we found several details about what new and/or required elements such business models might include. Primarily, we observed that "producing" valuable data (be it about the supply chain, operations, transactions, customers, or finance, or a combination) typically occurs in many places along the value chain.

Bharadwaj et al. (2013) call for transfunctional IT roles, new mandates for IT experts and CIOs, and generally higher familiarity among staff, but our results indicate a far more fundamental and structural incompability between traditional business models and business models for selling data. We found many challenges related to regarding data as sellable and monetizable private property in the first place and the need for creative arrangements to unlock the potential of data business. We believe that a wider need a research agenda for understanding requirements for data monetization, digital systemic innovations, and inter-firm data-based business models exists.

Our empirical evidence indicates that data-based business models mature along evolving paths even if data monetization could more radically disrupt business models. In selling and monetizing data, firms follow different paths, reach a different maturity level, and uncover different capabilities in providing data-monetization offerings. Few companies can find a data-monetization business model without navigating through several business model shifts. Typically, the path starts from an existing value chain position, existing data, and existing stakeholders, and some sort of external pressure instigates it. Companies then fairly quickly navigate through business model "pivots"—a resource-consuming endeavor that one can interpret as indicating their inability to disrupt the current business model. The following examples show how different companies we examined pivoted their data-monetization business models in practice.

6   Examples of Case-company Monetization Paths

6.1   Company E: Finance

Company E had more ongoing data-monetization projects than traditional finance companies. The company used its data mainly to improve internal decision making, management, and processes. The company had improved its rigid legacy IT systems to enable better flexibility and efficiency with a contemporary IT infrastructure. It had slowly begun changing from using data internally to using it externally (primarily to enhance the customer experience, not to obtain revenue). To do so, the company had shared its data with chosen third parties to provide discounts to end users. The company had also started to create a digital platform where external developers could use its data through APIs.

6.2   Company J: Healthcare

Healthcare company J wanted to make its internal decision making and processes more data driven. As a secondary goal, it wanted to improve the customer experience, yet it saw this goal more as a longer-term one. However, the company did monetize its data in its value network with insurance companies and pharmaceutical companies. This data exchange was a tradition in the industry, and it did not monetize data in newer ways. It acknowledged the potential ways it could use anonymous data but saw them as uncharted territory.

6.3   Company O: Manufacturing

Manufacturing company O, a pioneer in data monetization, had been optimizing its internal operations with data for a long time and nowadays provided data products and services to its customers. It offered raw data as datasets and a continuous service. It also provided more refined services through which
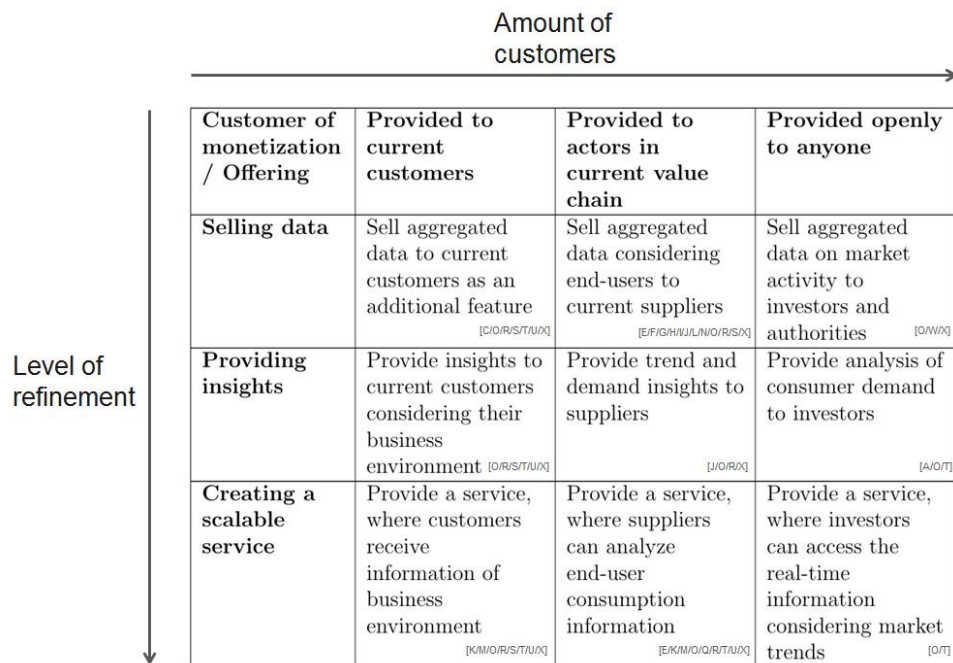
customers could receive insights through dashboards. For this manufacturing company, data was an integral part of their offering and tightly linked to their core products.

6.4 Company S: Media

Media company S used data in various ways as is common in the media industry. It monetized data in multiple ways in addition to internal data-based optimization and improved customer experience: by providing raw data to research companies, offering data-derived insights to advertisers, and building dashboards to a strategic partner. In addition, it sold raw data to third parties. In total, it provided data to current customers, other actors in its value network, and to third parties.

6.5 Company X: Transportation

Company X had valuable location data as a transportation company. The company mainly used this data to improve its internal processes and to optimize the customer experience, but it also provided data to other actors in its value network. It monetized data through continuous services to its partner network. In addition to revenue-creating services, the company had experimented with data-exchange pilots with a few chosen partners.

Amount of customers →

Customer of monetization / Offering	Provided to current customers	Provided to actors in current value chain	Provided openly to anyone
Selling data	Sell aggregated data to current customers as an additional feature [C/O/R/S/T/U/X]	Sell aggregated data considering end-users to current suppliers [E/F/G/H/I/J/L/N/O/R/S/X]	Sell aggregated data on market activity to investors and authorities [O/W/X]
Providing insights	Provide insights to current customers considering their business environment [O/R/S/T/U/X]	Provide trend and demand insights to suppliers [J/O/R/X]	Provide analysis of consumer demand to investors [A/O/T]
Creating a scalable service	Provide a service, where customers receive information of business environment [K/M/O/R/S/T/U/X]	Provide a service, where suppliers can analyze end-user consumption information [E/K/M/O/Q/R/T/U/X]	Provide a service, where investors can access the real-time information considering market trends [O/T]

Level of refinement ↓

Figure 1. Refined Business Model Alternatives for Data Monetization based on Empirical Information about Company Paths

The simplest and least innovative way to start monetizing data involves selling fairly unrefined data to one's current customers and partners. However, for many companies, changing from not sharing data to selling it to multiple companies represents a huge leap. Therefore, a first step to monetizing data involves offering it to one's customers and partners. Accordingly, the first data-monetization path involves demonstrating data's value to other actors. The second path involves providing value to a company's current customers and current value chain as it can more easily control offerings that it provides a limited number of partners. However, in the second path, a company provides more refined data (i.e., datasets or insights). Once companies have provided data to their current customers and the actors in their value chain, they follow the third path: openly offer it to any actor and, thereby, attract new customers from different industries. A company should follow this third path only after following the first two given the higher risks involved and the investments required. To conclude, it appears that companies need to go through each path to develop and evolve their existing business models because they adapt gradually to novel value offerings that depart substantially from old ones (Tikkanen, Lamberg, Parvinen, Kallunki, 2005). Our findings indicate that selling and monetizing data as a new business model for a company involves overcoming multiple organizational challenges simultaneously. In Figure 1, we illustrate our key

findings and develop the earlier business model categorization further by adding our empirical observations about paths. We show how the path logic details the concrete business model navigation options that companies have.

7   Managerial Implications

According to our results, most companies willingly move slowly in using and monetizing data either because they have a culture in which they see data as a tradable asset or because they have understandable concerns about privacy and security. A company will find it difficult to monetize data it it has not yet succeeded in exploiting it internally. As Najjar and Kettinger (2013) suggest, companies should acquire sufficient capabilities to address internal business needs before monetizing data: they could more easily achieve foreseeable and quicker results by investing in internal processes and current products rather than creating new data offerings.

McKinsey's often-cited white paper (Bradley & O'Toole, 2016) claims that 52 percent of Fortune 500 companies have gone bankrupt, been acquired, or ceased to exist due to business transformation. Such transformation has been most profound in industries where one can digitize assets, but we can expect the same phenomenon to happen in more traditional industries, which calls for renewed attention to generate revenue from data. However, if a company has created solid internal premises and use cases for data, monetizing data opens up new business opportunities.

According to our results, companies should not build data offerings built on unstable premises: they require a proper data infrastructure and an organizational culture that supports monetization. Our findings suggest that companies need to invest in their data infrastructure so that they can deliver high-quality and secure data that they can easily access and that they need capable employees who can refine, commercially exploit, and sell it. We identified a pattern that companies may most easily access potential customers for data monetization if they have a close business relationship with the selling company: current customers and companies in a company's value network will likely most readily find added value in the company's data. According to our data, strategic partnerships provide the easiest and the least risky route to monetizing data. However, given that many found it difficult to form value propositions from their data, we observe a need for companies to foster collaboration with other actors to better understand their data's value potential. We found explicit evidence that companies need to have or move into a favorable value chain position, be able to find a legitimate pricing logic, and think beyond current partners as soon as possible to further exploit data. We can also draw the conclusion that companies also need to train their employees to better understand data-based offerings' commercial aspects and to be able to operate in the interface between business and IT.

We also found that selling and monetizing data requires significant managerial support and strategic focus, even visionary leadership. We observed a pattern that many companies struggled to create viable business cases for data offerings, and it may be challenging to realize monetization's potential benefits with traditional tools. Thus, we found that companies need a more risk-tolerant strategy and to implement it with experimentation in mind, such as in parallel startups, corporate ventures, or collaboration with a startup. Companies establishing a separate subsidiary could resolve potential issues regarding the parent company's trust, core business, lack of agility and culture, and legislation in some cases. As a logical extension, establishing a separate business or acquiring another company involves less risk to a company's initial brand and trust to its data-monetization solutions. A separate organizational position could also leave room for experimentation given the lack of direct pressure to provide continuous results. The most successful initiative among the interviewed companies, in terms of direct short-term revenue and profit, was first launched as a general checkup and insight service. It then evolved into a separate internal venture that established a business unit that provided a scalable service. The scalable service provided business process outsourcing customers with the ability analyze every, or any, customer in their go/no-go decisions.

As a general inference from our results, we suggest several steps to monetize data that companies could take regardless of which business model they choose (see below and Figure 2). These steps represent organizational decisions and tasks in the selling company that we found in the most successful data-monetization cases among the interviewed companies. These companies were able to create a continuous system to monetize data rather than operating on a one-off, project basis. The steps include:

- **Exploit a central position in the value chain to collect data:** a company could use its position at the intersection of a value chain to monetize data by collecting it from one side and delivering it to the other. For example, a company's suppliers might value information from its customers, or its customers might value data on other customers.

- **Appoint a champion**: internal challenges and needed arrangements vary by company, but different champion roles seem to fix these factors regardless of the champion's background.

- **Discuss with multiple different partners**: companies could talk to other companies to better understand their data's potential in terms of business opportunities. They may find suitable discussion partners in their current value chain or network but should also venture outside given that a common end customer may be enough to add value to the data.

- **Ensure good-quality and accessible data:** given that data monetization's value depends on the data, companies should have adequate quality data and be able to easily access it. Companies need a proper data infrastructure to monetize data in an ongoing manner.

- **Establish an independent team, subsidiary or company:** using a subsidiary or a distinct team may result in more flexible experimentation. It makes agile development easier: the company's culture does not slow down the process, and experimentation through a new brand does not adversely affect the position of or trust in the initial brand.

- **Validate data value with a chosen partner**: companies need to assess data's and data offerings' value before creating a comprehensive monetization plan. For example, companies could set up a pilot project. Companies can use data from the pilot when pricing and communicating value proposition in the future.

- **Create processes that support continuous monetization**: establishing roles and processes to support ongoing development could help companies ensure they monetize data in a cost-effective and high-quality manner.
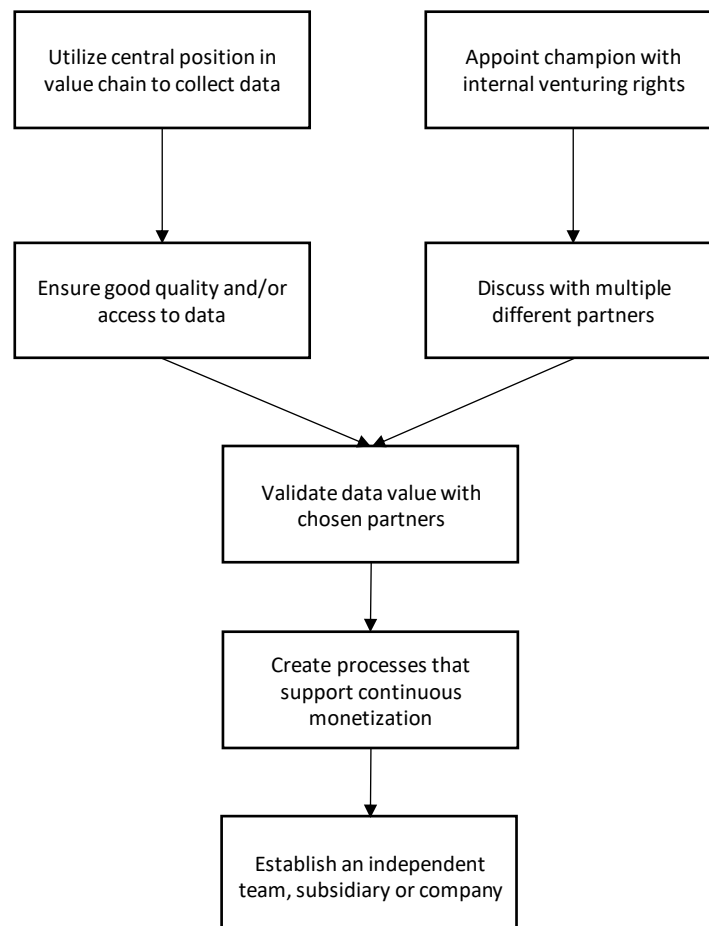
Figure 2. Recommended Steps on the Path to Data Monetization

These steps may be useful in achieving data monetization in practice. However, as we observe in Section 4, companies cannot simply integrate data monetization into their current offerings, and how they monetize data depends on the strategy they choose. Therefore, they should choose a suitable data offering and an appropriate customer basis. Note also that the way companies present data to their partners and customers can be a great asset for their brand if done well, but they can produce disastrous results if they sell and repurpose data in a way that customers despise. As a good example, Google has made its quite boring search box into a place for people to go to search for information. At the other end of the spectrum, Cambridge Analytica not only destroyed its own reputation but also harmed its data provider, Facebook, and laid bare some questionable practices.

8   Limitations and Future Research

As with any study, this study has its limitations. The research interviews reflect the opinions of the interviewees, who are prone to subject error and bias. Indeed, given that only one or few interviewees represented each company, we cannot comprehensively understand each company's respective standpoints on data monetization. The data rather represents the interviewees' positions and viewpoints. Moreover, even though we guaranteed anonymity and confidentiality, the interviewees may have wanted to hide or downplay their organizations' real plans. As the data showed, many companies do not wish to discuss data monetization publicly given that society tends to see it as a dubious act. As Alvesson and Sköldberg (2018, p. 354) also note, interviewees may exaggerate their personal achievements or deliberately present themselves in a rational light. To evaluate whether they did so, we conducted expert interviews after our primary interviews with senior IT consultants that had worked with many of the firms in our sample. These supplementary interviews did not expose any cases of misrepresentation.

Qualitative case studies commonly suffer from weakly transferable results, and we acknowledge that our results explain data-monetization practices in a specific context (Miles & Hubermann, 1994; Yin, 2009). Our findings may not fully generalize to other companies. Therefore, we suggest that, in the main, our results apply to companies that operate in the EU given they all operated under common legislation on data privacy and to large, professional companies in Western countries that operate under similar working cultures and ethics as the ones we studied. Nevertheless, given that the EU's data-privacy legislation functions as a global moral guideline and benchmark (Solon, 2018b). Moreover, our case companies represented a wide range of industries and product/service types, and the interviewees held various relevant job titles and positions, which we believe further alleviates some generalizability problems.

We clearly need more research on data monetization, particularly given that companies have shown an increased interest in it. As interest in monetizing data continues to grow and increasing computing capacity and evolving management norms generate more potential for growth, cutting-edge research may allow firms to make fewer expensive mistakes.

We can see the process model that we present in Figure 2 as the starting point for an initial process theory about how companies that that instantiate the suggested steps will do better in monetization. We need further experience and research to detail better the ways to perform the activities in each step.

We also need research that accommodates companies' aspirations and both consumer and societal restrictions. For example, we need research on the line between acceptable and unacceptable data-monetization practices and how companies can alleviate data-monetization concerns and create new innovative data offerings in a morally sound way. Companies and academics alike need to monitor how data monetization evolves in terms of legislative changes, improvements in computational methods and capacity, and for the social system we live in. Future research should provide guidelines that prevent customers from becoming inadvertent victims of data misuse and pay attention to data security and privacy research and practice.

9   Conclusion

In this paper, we discuss how different organizations can capture, value, and monetize the data assets that they possess. Little research has examined the components of data-based business models that deal with capturing data value and the challenges involved in monetizing data to date. Therefore, we call for more research in this area. As companies move from selling products to selling or renting services, the value of data increases exponentially.

We used cases to highlight problems and the best practices of monetizing companies' data assets. Furthermore, we classified case company business models from a monetization point of view by identifying business potential through a matrix of customer and offering types. We then identified data-monetization paths and, based on them, provide a preliminary model for testing data monetization in companies. The results should offer possibilities for further research and provide also practical tools for companies that are entering the field.

Researchers can conduct further research in several obvious areas. As data-monetization research remains immature, we need more in-depth case studies that examine companies transforming their offerings through data sales. Since researchers have speculated about data-monetization value chains but still largely unexplored them, it would be of interest to understand the role that data aggregators and refiners play in data monetization, how they create value, and how different parties can capture it. Currently, most analyses and cases deal with selling customer data or monetizing customer behavior to marketers.

Acknowledgements

References

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis, 16*(1), 3-9.

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly, 25*(1), 107-136.

Alvesson, M., & Sköldberg, K. (2018). *Reflexive methodology: New vistas for qualitative research* (3rd ed.) Thousand Oaks, CA: Sage.

Amit, R., & Zott, C. (2012). Creating value through business model innovation. *MIT Sloan Management Review, 53*(3), 41-49.

Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). Transformational issues of big data and analytics in networked business. *MIS Quarterly, 40*(4), 807-818.

Barclays. (2016). *Barclays helps SMEs access "big data" to aid growth.* Retrieved from https://newsroom.barclays.com/r/3336/barclays_helps_smes_access__big_data__to_aid_growth

Bean, R. (2017). How companies say they're using big data. *Harvard Business Review.* Retrieved from https://hbr.org/2017/04/how-companies-say-theyre-using-big-data

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: Toward a next generation of insights. *MIS Quarterly*, *37*(2), 471-482.

Birkinshaw, J., Brannen, M. Y., & Tung, R. L. (2011). From a distance and generalizable to up close and grounded: Reclaiming a place for qualitative methods in international business research. *Journal of International Business Studies*, *42*(5), 573-581.

Bradley, C., & O'Toole, C. (2016). An incumbent's guide to digital disruption. *McKinsey Quarterly*, *52*(3), 76-85.

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of "big data"? *McKinsey Quarterly*. Retrieved from https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/are-you-ready-for-the-era-of-big-data.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188.

Choo, C. W. (1996). The knowing organization: How organizations use information to construct meaning, create knowledge and make decisions. *International Journal of Information Management*, 16(5), 329-340.

Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Platforms and infrastructures in the digital age. *Information Systems Research*, *29*(2), 381-400.

Crain, M. (2018). The limits of transparency: Data brokers and commodification. *New Media & Society*, *20*(1), 88-104.

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

Debatin, B., Lovejoy, J. P., Horn, A. K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, *15*(1), 83-108.

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359-363.

Dillet, R. (2014). Nest uses its data to turn electric utilities into cash cows. *TechCrunch*. Retrieved from https://techcrunch.com/2014/04/18/nest-uses-its-data-to-turn-electric-utilities-into-cash-cows/

European Commission. (2015). *European Parliament adopts European Commission proposal to create safer and more innovative European payments.* Retrieved from http://europa.eu/rapid/press-release_IP-15-5792_en.htm?locale=en

European Commission (2018). *Communication from the commission to the European Parliament and the Council.* Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1517578296944&uri=CELEX%3A52018DC0043

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research. *Organizational Research Methods*, *16*(1), 15-31.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, *8*(4), 597-606.

Haddara, M., & Elragal, A. (2015). The readiness of ERP systems for the factory of the future. Pr*ocedia Computer Science*, *64*, 721-728.

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

Jensen, C., Potts, C., & Jensen, C. (2005). Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies*, *63*(1-2), 203-227.

Lakomaa, E., & Kallberg, J. (2013). Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. *Access IEEE*, *1*, 558-563.

Lewis, A., & McKone, D. (2016). To get more value from your data, sell it. *Harvard Business Review*. Retrieved from https://hbr.org/2016/10/to-get-more-value-from-your-data-sell-it

Liew, A. (2007). Understanding data, information, knowledge and their inter-relationships. *Journal of Knowledge Management Practice*, *8*(2), 1-16.

Lindman, J., Kinnari, T., & Rossi, M. (2016). Business roles in the emerging open-data ecosystem. *IEEE Software*, *33*(5), 54-59.

Matsakis, L. (2018). Facebook's targeted ads are more complex than it lets on. *Wired*. Retrieved from https://www.wired.com/story/facebooks-targeted-ads-are-more-complex-than-it-lets-on/

McDermott, C. M., & O'Connor, G. C. (2002). Managing radical innovation: An overview of emergent strategy issues. *Journal of Product Innovation Management*, *19*(6), 424-438.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Thousand Oaks, CA: Sage.

Opresnik, D., & Taisch, M. (2015). The value of big data in servitization. *International Journal of Production Economics*, *165*(), 174-184.

Najjar, M. S., & Kettinger, W. J. (2013). Data monetization: Lessons from a retailer's journey. *MIS Quarterly Executive*, *12*(4), 213-225.

Negash, S. (2004). Business intelligence. *Communications of the Association for Information Systems*, *13*, 177-195.

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

Osterwalder, A. Pigneur, Y., & Tucci, C (2005). Clarifying business models: Origins, present, and future of the concept. *Communications of the Association for Information Systems*, *16*, 1-25.

Pantelis, K., & Leiponen, A. (2013). Understanding the value of (big) data. In *Proceedings of the IEEE International Conference on Big Data* (pp. 38-42).

Parmar, R., Mackenzie, I., Cohn, D., & Gann, D. (2014). The new patterns of innovation. *Harvard Business Review.* Retrieved from https://hbr.org/2014/01/the-new-patterns-of-innovation

PatientsLikeMe. (2016). *How does PatientsLikeMe make money?* Retrieved from https://support.patientslikeme.com/hc/en-us/articles/201245750-How-does-PatientsLikeMe-make-money-

Piccinini, E., Hanelt, A., Gregory, R., & Kolbe, L. (2015). Transforming industrial business: The impact of digital transformation on automotive organizations. In *Proceedings of the 36th International Conference on Information Systems.*

Quah, D. (2003). *Digital goods and the new economy* (NoCEPR Discussion Paper No. 3846). Retrieved from https://ssrn.com/abstract=410604

Rathore, M. M., Ahmad, A., Paul, A., & Rho, S. (2016). Urban planning and building smart cities based on the Internet of things using big data analytics. *Computer Networks*, *101*, 63-80.

Rayna, T. (2008). Understanding the challenges of the digital economy: The nature of digital goods. *Communications & Strategies*, 71, 13-16.

Reeves, S., Albert, M., Kuper, A., & Hodges, B. D. (2008). Why use theories in qualitative research? *BMJ*, *337*, a949.

Ritchie, J., & Lewis, J. (2014). Qualitative research practice: A guide for social science students and researchers. Thousand Oaks, CA: Sage.

Rossman, J. (2016). *The Amazon way on IoT: 10 principles for every leader from the world's leading Internet of things strategies.* Bellevue, WA: Clyde Hill Publishing.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* (vol. 5). New York, NY: Pearson.

Saxe, R., & Weitz, B. A. (1982). The SOCO scale: A measure of the customer orientation of salespeople. *Journal of marketing research*, *19*(3), 343-351.

Schüritz, R., & Satzger, G. (2016). Patterns of data-infused business model innovation. In *Proceedings of the 18th IEE Conference on Business Informatics.*

Sharma, S., & Gupta, B. (2018). Information privacy on online social networks: Illusion-in-progress in the age of big data? In A. V. Deokar, A. Gupta, L. S. Iyer, & M. C. Jones (Eds.), *Analytics and data science* (pp. 179-196). Berlin: Springer.

Sheehan, K. B. (2002). Toward a typology of Internet users and online privacy concerns. *The Information Society*, *18*(1), 21-32.

Solon, O. (2018a). Facebook says Cambridge Analytica may have gained 37m more users' data. *The Guardian.* Retrieved from https://www.theguardian.com/technology/2018/apr/04/facebook-cambridge-analytica-user-data-latest-more-than-thought

Solon, O. (2018b). How Europe's "breakthrough" privacy law takes on Facebook and Google. *The Guardian.* Retrieved from https://www.theguardian.com/technology/2018/apr/19/gdpr-facebook-google-amazon-data-privacy-regulation

Spijker, A. V. (2014). *The new oil: Using innovative business models to turn data into profit.* Basking Ridge, NJ: Technics Publications.

Stringer, R. (2000). How to manage radical innovation. *California Management Review*, *42*(4), 70-88.

Tamro (2017a). *Solutions for pharmacies.* Retrieved from http://www.tamro.fi/en/SolutionsAndProducts/ForPharmacies/Pages/Solutions-for-pharmacies.aspx

Tamro. (2017b). *Tamro HANDY—aina valmiina tulokselliseen asiakaskohtaamiseen.* Retrieved from http://www.tamro.fi/fi/PalvelutJaTuoteryhmat/L%C3%A4%C3%A4ke-%20ja%20terveystuoteyrityksille/Laakeyrityksille/Sivut/Tamro-HANDY.aspx

Teece, D. J., & Linden, G. (2017). Business models, value capture, and the digital enterprise. *Journal of Organization Design*, *6*(1), 1-14.

Terho, h., Haas, A., Eggert, A., & Ulaga, W. (2012). "It's almost like taking the sales out of selling"—towards a conceptualization of value-based selling in business markets. *Industrial Marketing Management*, *41*(1), 174-185.

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

Tikkanen, H., Lamberg, J. A., Parvinen, P., & Kallunki, J. P. (2005). Managerial cognition, action and the business model of the firm. *Management Decision*, *43*(6), 789-809.

Valentino-DeVries, J. (2018). Service meant to monitor inmates' calls could track you, too. *The New York Times.* Retrieved from https://www.nytimes.com/2018/05/10/technology/cellphone-tracking-law-enforcement.html

Walker, S. J. (2014). Big data: A revolution that will transform how we live, work, and think. *International Journal of Advertising*, *33*(1), 181-183.

Whitmore, T. (2016). Data monetization: Who's doing it right (now). *Blue Hill Research*. Retrieved from http://bluehillresearch.com/data-monetization-whos-doing-it-right-now/

Wiesche, M., Jurisch, M. C., Yetton, P. W., & Krcmar, H. (2017). Grounded theory methodology in information systems research. *MIS Quarterly*, *41*(3), 685-701.

Wixom, B. H., & Ross, J. W. (2017). How to monetize your data. *MIT Sloan Management Review*, *58*(3), 10-13.

Woerner, S. L., & Wixom, B. H. (2015). Big data: extending the business strategy toolbox. *Journal of Information Technology*, *30*(1), 60-62.

Yin, R. K. (2009). *Case study: Design and methods.* Thousand Oaks, CA: Sage.

Yousif, M. (2015). The rise of data capital. *IEEE Cloud Computing*, *2*(2), 1-11.

Zott, C., Amit, R., & Massa, L. (2011). The business model: Recent developments and future research. *Journal of Management*, *37*(4), 1019-1042.

About the Authors

Petri Parvinen is Professor of Strategic Marketing and Management at Helsinki University, Finland and part-time Professor of Selling and Sales Management at Aalto University, Finland. He leads several interdisclinary groups of scientists studying the benefits of new technology in sales, marketing and education. He serves business and society actively by popularizing his research that has appeared in journals such as *Journal of Product Innovation Management*, *Value in Health*, *Journal of Management Studies*, *Journal of Statistical Software*, *Industrial Marketing Management,* and *International Journal of Electronic Commerce.*

Essi Pöyry is Post-Doctoral Researcher at Centre for Consumer Society Research at University of Helsinki, Finland. Her research concentrates on the impact of digital and social media on customer and citizen behavior. Her research has been published in *Computers in Human Behavior, International Journal of Electronic Commerce*, *Electronic Commerce Research and Applications*, *International Journal of Strategic Communication*, *Journal of Marketing Management*, *Journal of Personal Selling & Sales Management*, among others.

Robin Gustafsson is Associate Professor of Strategic Management at the Department of Industrial Engineering and Management at Aalto University. His research focus is on strategy, organization, and policy in technology induced industry and market disruptions. His recent research explores how digital is disrupting existing industries and markets, platform economy, platform design and strategy, sources of competitive advantage, data monetization, and agile corporate digital strategy and transformation work. His research has been published in top tier journals, such as the *Academy of Management Journal*, *International Journal of Management Reviews*, *Information Systems Research*, *Research Policy*, *Technological Forecasting and Social Change*, and *Organization Studies*. He is an expert on flipped classroom and experienced-based learning methods, including the Harvard case teaching method, and LEGO Serious Play facilitation method.

Miikka Laitila is a Business Designer at Futurice with an interest in novel data-driven business opportunities. His expertise lies in data-driven business, platform opportunities and digital strategy. As a business designer, he approaches new opportunities through customer-centric, strategic and data lenses often using design thinking, lean startup and agile methods. In recent years, he has gained experience in projects for financial, pharmaceutical, logistics and manufacturing industries and holds MSc (Tech.) at Aalto University. His master thesis was among the first in the world to systematically explore data monetization and data-based business models.

Matti Rossi is Professor of Information Systems at Aalto University School of Business. He is a past president of the Association for Information Systems, and has been the principal investigator in several major research projects funded by the Technological Development Center of Finland and the Academy of Finland. He was the winner of the 2013 Millennium Distinction Award of the Technology Academy of Finland for open-source and data research. His research papers have appeared in journals such as *MIS Quarterly*, the *Journal of the AIS*, *Information and Management* and *Information Systems*. He has been a senior editor of the *Journal of the AIS* and *Database*, and an associate editor for *MIS Quarterly*, and he is the past editor in chief of *Communications of the Association for Information Systems*.