Aalto University
School of Science
Degree Programme in Information Networks

Miikka Laitila

Data monetization:
Utilizing data as an asset to generate new revenues for firms

Master's Thesis
Espoo, October 3, 2017

Supervisor: Professor Robin Gustafsson
Advisor: Mika Ruokonen D.Sc. (Econ)

Aalto University
School of Science
Degree Programme in Information Networks

ABSTRACT OF
MASTER'S THESIS

Author: Miikka Laitila

Title:
Data monetization: Utilizing data as an asset to generate new revenues for firms

Date: 03.10.2017 **Pages:** 100

Major: Information Networks **Code:** SCI3047

Supervisor: Professor Robin Gustafsson

Advisor: Mika Ruokonen D.Sc. (Econ)

The discussion around big data and the potential use of it has emerged rapidly during recent years. The increasing data masses with development in IoT, cloud computing, analytics and legislation enable new possibilities to utilize the collected data. The literature has focused mainly on how companies can use their data to optimize current processes and decision-making or how to improve current services and products. Yet, companies can utilize the collected data in new ways: They can monetize the collected data by creating new data-based services or by selling it. Nevertheless, the level and popularity of data monetization for different companies has stayed unknown.

This thesis focuses on the concept of data monetization: What is data monetization, in which ways data can be monetized and which factors affect the monetization. As data monetization has not been studied widely, this thesis provides a distinct definition for the nascent term. The literature review of this thesis forms a structured summary considering the concept of data monetization and different ways how to monetize data. The barriers preventing data monetization are discussed as well. The empirical part of this thesis examines the state of monetization in big Finnish companies. The inductive case study of 19 different companies is carried out through interviews.

The results of this thesis indicate data monetization as one potential way to utilize current data, yet rather rare and niche business for majority of companies. The variety and combination of different barriers slow down the creation of monetization solutions. The results suggest that companies can benefit from monetization in multiple ways, especially by improving current offerings and relationships with customers and partners through monetization. This thesis provides a framework to interpret the phenomenon of data monetization as well as different ways to monetize data. The framework provides a novel basis for future research to study the phenomenon further. In addition, concrete recommendations help companies and other actors to evaluate how they should approach data monetization.

Keywords: Data monetization, data commercialization, new revenue, value of data, data reuse, strategy, business model

Language: English

Aalto-yliopisto
Perustieteiden korkeakoulu
Informaatioverkostojen koulutusohjelma

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Miikka Laitila

Työn nimi:
Datan monetisaatio: Data yrityksien uuden liikevaihdon pohjana

Päiväys:	03.10.2017	**Sivumäärä:**	100
Pääaine:	Informaatioverkostot	**Koodi:**	SCI3047
Valvoja:	Professori Robin Gustafsson		
Ohjaaja:	Mika Ruokonen TkT		

Keskustelu ”big datasta”ja sen mahdollisesta hyödyntämisestä on lisääntynyt viime vuosina. Yritysten valtavat datamäärät ja kehitys esineiden internetissä, pilviratkaisuissa ja analytiikassa sekä yleinen lainsäädännön muuttuminen mahdollistavat uudet tavat hyödyntää kerättyä dataa. Alan kirjallisuus on keskittynyt yrityksien mahdollisuuksiin käyttää dataa sisäisten prosessejen ja päätöksenteon tehostamiseen tai nykyisten tuotteiden ja palveluiden parantamiseen. Yritykset voivat kuitenkin hyödyntää dataa myös uusilla tavoilla: Datalla voi luoda uusia datapohjaisia palveluja tai sitä voi myydä toisille yrityksille. Tällaista datan hyödyntämistä uuden liiketoiminnan lähteenä ei ole tutkittu laajasti, eikä monetisoinnin yleisyys yrityksissä ole tiedossa.

Tämä tutkimus keskittyy datan monetisointiin: Mitä on datan monetisointi, millä tavoin dataa voidaan monetisoida ja mitkä asiat vaikuttavat datan monetisointiin. Koska datan monetisointia ei ole tutkittu laajasti, tämä tutkimus tarjoaa selkeän määritelmän ilmiölle. Tämän tutkimuksen kirjallisuuskatsaus muodostaa jäsennellyn koosteen datan monetisoinnista ja eri tavoista kaupallistaa dataa. Katsaus käsittelee myös datan monetisointia estäviä tekijöitä yrityksissä. Tutkimuksen empiirinen osuus tarkastelee datan monetisoinnin tilaa suurissa suomalaisissa yrityksissä. Induktiivinen 19 yrityksen tapaustutkimus toteutetaan haastatteluilla.

Tutkimuksen tulokset osoittavat, että datan monetisointi on yrityksille yksi tapa hyödyntää dataa, vaikkakin vielä melko harvinainen ja kapea liiketoimintamahdollisuus. Datan monetisoinnin yleistymistä hidastaa erilaisten esteiden suuri määrä sekä niiden yhteisvaikutus. Tutkimuksen tulosten perusteella yritykset voivat hyötyä monetisaatiosta esimerkiksi parantamalla nykyisiä suhteita asiakkaiden ja partnereiden kanssa datan monetisoinnin kautta. Tämä työ esittelee viitekehyksen, joka auttaa ymmärtämään datan monetisointia ja sen eri toimintatapoja. Viitekehys luo uutta pohjaa jatkotutkimuksille datan monetisaatiosta. Lisäksi työ tarjoaa konkreettisia suosituksia, jotka auttavat yrityksiä ja toimijoita arvioimaan parhaita tapoja lähestyä datan monetisointia.

Asiasanat:	Datan monetisointi, datan kaupallistaminen, datan arvo, datan uudelleenkäyttö, strategia, liiketoimintamalli
Kieli:	Englanti

Acknowledgements

As with any project, this would not have been possible without great people supporting me. Thanks everybody, it is truly amazing to be surrounded with such a wonderful people.

First of all, I'd like to thank my supervisor Robin Gustafsson. Thanks for guiding me during the research process and providing me insights as I learnt more and more about data monetization. Thanks to you, the ideas and findings are much more structured, refined and explicit.

Thanks all wonderful people at Futurice for inspiring discussions and ideas. Thanks especially to my advisor Mika Ruokonen for giving me vital feedback on drafts and guiding me to right direction with a broad topic. Thanks to Tuomas Syrjänen for initially coming up with this topical idea. It has been a pleasure to dive into the novel field of monetization.

I want to thank Information Networks' guild Athene, for introducing me to great and inspiring people. It is much easier to get by with a help from friends.

I want to thank my parents for supporting me during my life and studies. Thanks for letting me to follow my own path, even though it less surprisingly took me to Otaniemi. Thanks to my brothers, Nuutti for being the greatest little brother one can have and Tuukka for showing the way with his example.

And lastly, thank you Heta, for everything.

Helsinki, 03.10.2017

Miikka Laitila

Contents

Abbreviations and Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
B2B	Business to Business
B2C	Business to Consumer
CDO	Chief Digital Officer
CIO	Chief Information Officer
Fintech	Financial technology
GDPR	General Data Protection Regulation
IoT	Internet of Things
PSD2	Payment Service Directive 2

Chapter 1

Introduction

This thesis studies the role of data monetization for Finnish companies: How companies can utilize their current data assets to generate new revenue? This study consists of theoretical and empirical parts focusing on data monetization: How companies can utilize and monetize their data and what kind of issues affect them? The theoretical part consists of literature review covering the concept of data monetization and practices and issues affecting it. The empirical part of this research is conducted by interviewing Finnish companies from different industries.

This first chapter acts as an introduction to the research. First the background and motivation of the thesis is presented. Then the research problem is defined as well as three different research questions presented. The goal and the scope of the study are discussed, and finally the structure of the study is introduced.

1.1 Background and motivation

The data possessed by companies increase continuously (Brown et al., 2011; Yousif, 2015; Manyika et al., 2011). As there is more data available, the amount of possibilities to utilize it broader increases as well (Opresnik and Taisch, 2015; Chen et al., 2012; Mayer-Schönberger and Cukier, 2014; Choo, 1996). While new ventures and multinational digital companies disrupt traditional businesses, the pressure encourages companies to utilize their data more broadly (Spijker, 2014; Yousif, 2015). This can be seen in industries such as transportation, accommodation, media and retail, where companies like Uber, Airbnb, Facebook, Netflix and Amazon have transformed the competition with data-driven practices and offerings (Mackenzie et al., 2014; Yousif, 2015; Spijker, 2014; Rossman, 2016). As different companies have

more and more data, the temptation and possibilities to utilize it increase.

Companies can use data in many different ways, for example to understand the changing business environment, create new innovations as well as to guide decision making (Choo, 1996). Still, the utilization of data has focused on the internal data usage and improving current products and services with data (Bean, 2017; Wixom and Ross, 2017; Thomas and Leiponen, 2016). Using data to create new products and services, or even selling the data or analysis derived from it, have been studied scarcely in business and academic literature (Thomas and Leiponen, 2016; Fred, 2017; Najjar and Kettinger, 2013). In academic literature, the phenomenon of big data and the utilization of it have been popular topics over the last few years (Thomas and Leiponen, 2016; Mayer-Schönberger and Cukier, 2014; Brown et al., 2011; Manyika et al., 2011). Due to popularity of digitalization and big data, the interest in the topic of data monetization has grown as well (Fred, 2017). The idea of monetization, creating new value and revenue from data, is not new (Thomas and Leiponen, 2016), yet the concept has not been studied properly (Thomas and Leiponen, 2016; Fred, 2017; Najjar and Kettinger, 2013). The changing business environment creates new possibilities for data utilization, data monetization as being a one option for a broader data usage (Mackenzie et al., 2014; Opresnik and Taisch, 2015). This has created a research gap, as companies are increasingly using the data in new ways, while the academic literature does not cover these new data derived offerings.

This thesis explores the concept and nature of data monetization, as well as the state of data monetization in big Finnish companies. Global digital pioneers, such as Google and Amazon have been appraised as data-driven companies (Spijker, 2014; Rossman, 2016; Yousif, 2015), but the levels of data utilization and monetization in Finnish companies are unknown: The data is increasingly the most important asset companies possess (Yousif, 2015; Spijker, 2014), but how broadly companies utilize it in practice? Do companies utilize their data by creating new services based on data or do they monetize their data directly by selling it or providing premium analysis based on data? If companies do not have these kind of monetization solutions, what kind of issues then affect the monetization and prevent a broader utilization of company's data? By understanding better the monetization, its practices and the different issues preventing it, companies can easier engage in monetization activities and learn to utilize their data assets better. The improved understanding of data monetization contributes to different fields of research, as big data and its utilization, new business models, multi sided platforms, legislation and naturally strategical and technical issues related to data, are all interconnected to data monetization. The better comprehension of current business environment related to data makes it easier for compa-

nies, researchers and other stakeholders to address the issues and develop the knowledge of monetization further.

1.2 Research problem

The aforementioned research context lays a premise for the research problem. The identified research gap is approached with the following research problem:

What does data monetization mean for companies today?

This thesis aims to understand better the phenomenon of data monetization; what is it, how it can be practiced and what affects it. The approach to the problem is strategical: The goal is to understand the role of monetization for companies and different reasonings to practice or not to practice monetization. Thus, the research problem can be further divided to three different research questions:

RQ1: What does data monetization entail for firms today?

RQ2: How do firms monetize their data?

RQ3: What affects data monetization done by firms?

The RQ1 aims to create a sufficient basis to answer RQ2 and RQ3. All of the presented research questions are first approached with literature review in *Chapter 2*. Yet, RQ1 creates a premise to understand the more profound research questions RQ2 and RQ3. As the data monetization is a nascent concept in business and academic literature, a sufficient understanding of the concept, practices and different issues are required before empirical research.

1.3 Goal and scope of the study

The objective of this thesis is to provide a thorough understanding of data monetization. The concept of data monetization will be explained, defined and presented; what it is, how companies can monetize their data, what they are doing currently and which things affect the monetization?

The theoretical goal of this thesis is to provide concrete information on the topic of data monetization: The academic literature of data monetization is scarce, as acknowledged by Fred (2017), Najjar and Kettinger (2013) and Thomas and Leiponen (2016). The goal is to provide a comprehensive definition for the term, as currently there is no clear definition of data

monetization (Fred, 2017; Thomas and Leiponen, 2016). In addition to elaborating the phenomenon, the goal is to clarify the nature and content of monetization. For academic research, this thesis aims to provide practical examples and theoretical groundwork for data monetization. This thesis aims to summarize the associated academic literature considering data monetization, while contributing with empirical results to support the theory creation around data monetization. The aim is to provide more suitable tools and and frameworks to approach the concept of data monetization.

The practical objective of this thesis is to provide knowledge regarding data monetization in Finnish industries. This thesis aims to establish a coherent understanding of data monetization, its practices and the issues enabling and preventing it. The goal is to understand how big Finnish companies monetize their data. In addition, the focus is on the barriers preventing the potential monetization: Which issues companies struggle with in their data initiatives? By understanding the current state of monetization and gravity of barriers, companies as well as other stakeholders can develop easier the practices of data monetization further. Yet, the practical goal of this thesis is to improve the knowledge on data monetization: What does it contain, what does it mean for companies and how companies could approach it?

Since the focus area of this thesis is broad, and the data monetization has not been studied widely, it is necessary to set limitations related to the scope of the study. This research is performed as a qualitative study, data collection done by interviews. The approach to the data monetization is strategical, as there are a lot of technical issues and practicalities considering monetization. Different technical details related to data, for example data bases and analytics are excluded from this thesis. Thus, the focus is on strategical decisions regarding monetization: Do companies monetize their data? If they do not, why this is the case? The monetization of data is limited in this study to consider monetization, where other actors, such as customers, partners or other companies, are involved. Therefore, if data is used internally to reduce costs or optimize current processes, it is not considered as data monetization, even though some literature include it in the definition of data monetization.

The interviewees are aimed to be people working as CDOs, CIOs, Directors or in similar roles, where people have a strategical view on data initiatives. As there is not much information considering successful monetization practices, the goal is to cover multiple industries, while focusing on Finnish big companies. This affects the sample size, as it is required to be sufficient enough to cover multiple industries, at least 15 companies. By limiting the amount of industries, the study could provide more compact results, yet the understanding of current monetization in Finland would be limited to few

industries. Therefore, different industries are not restricted, as the focus is on big, primarily publicly listed, companies and their potentially valuable data.

1.4 Structure of the study

This thesis is divided into five different parts:

1. Introduction
2. Theoretical background
3. Methodology
4. Results
5. Discussion

The first chapter describes the background and motivation for the study, the research problem and research questions as well as the objectives and the scope of the study.

The second chapter of this thesis is a literature review, where relevant academic literature is presented to understand the phenomenon more thoroughly. The focus of the theoretical background is on the concept of data as an asset, drivers for increasing data usage and data monetization. Especially data monetization, practices of monetization as well as different issues affecting monetization are presented to establish a basis for empirical research.

Following the theoretical background, the third chapter presents the methodology used in this thesis. The research setting, methodology and data collection and analysis are discussed and presented thoroughly.

The fourth chapter presents the results from the empirical part of the study. The structure is similar to one presented in *Theoretical background*, as the value of data and changing business environment are discussed first, before focusing on current data monetization and barriers preventing the potential monetization.

After presenting the results of the study, the discussion is provided. The fifth chapter concludes answers to research questions and discusses the findings of the study. Additionally, a framework considering ways to monetize data and concrete recommendation are provided to ease the approach to data monetization. Practical and theoretical implications are provided, in addition to critical discussion considering limitations of the study. Lastly a conclusion is provided to summarize the thesis.

Chapter 2

Theoretical background

This chapter consists of the theoretical background for data monetization. Literature review focuses on data and how it can be regarded as an asset, on data monetization and its business models as well as common issues affecting monetization and utilization of data. The aim of theoretical background is to provide a sufficient basis for empirical research, as well as to bind this thesis into current academic literature. Firstly, data is defined, the basis for value of data is explained, as well as the changing business environment around data is presented. After explaining the concept of data, the phenomenon of data monetization is defined, different business models presented as well as different issues affecting monetization. In the end of the chapter a short summary is provided to conclude the findings from literature review.

2.1 Data as an asset

This section presents theoretical background for data, and how it can be considered as an asset. The definition of data used in this research is defined, as well as the value of data is discussed.

2.1.1 Definition of data

To understand data and its value, it's important to define data. Data, information and knowledge are often used in the same context, but their definitions slightly vary depending on the author (Liew, 2007; Leidner and Alavi, 2001; Ackoff, 1989). Data is often considered as a set of symbols that represent the properties of events and objects (Ackoff, 1989). For Liew (2007), data is unrefined and unfiltered information or a set of discrete objective facts about events. Leidner and Alavi (2001) define data as facts, raw num-

bers, whereas information is processed and interpreted data and knowledge is personalized information. Knowledge is often considered as actionable information or meaningful links between information and its applications (Liew, 2007; Ackoff, 1989). Regardless of the right definition, data can be seen as the raw material, which companies derive from their activities (Yousif, 2015). This raw data can be refined further into information or knowledge, which can hold more value to its owner and users of data. Therefore, it is seen as a raw material, whereas information and knowledge as refinements derived from data.

Yousif (2015) describes, that data is non-rivalrous, non-fungible and an experience good. Non-rivalrous means that one single data set can be exploited by multiple actors at the same time, whereas for physical products, for example a single car or a liter of paint can be used by only one actor at once. Non-fungibility relates to the variety of data; Different data sets contain different information and hold different value. Therefore, one cannot trade one data for another, as one can trade one dollar for another dollar, since data sets are not similar. Experience good relates to the value of data, since the value is realized only after one has used the data. Without utilizing the data, it is not possible to define its value (Yousif, 2015; Koutroumpis and Leiponen, 2013). This sets an interesting basis for data valuation and eventually monetization, since the real value of data is initially unknown.

In this thesis, data is considered *unrefined set of discrete objective facts*. Data has different features, such as being non-rivalrous, non-fungible and an experience good. Information and knowledge are seen as refinements of data.

2.1.2 Value of data

Data can be considered as raw material for companies to create new kind of value (Yousif, 2015). Value as a term relates to the usefulness, quality, importance, price and worth of a subject (Fred, 2017; Zeithaml, 1988). Yet, technically high quality data is not automatically valuable (Fred, 2017). As Fred (2017) notes, value is relative to individual's perception and context, therefore it is subjective. Koutroumpis and Leiponen (2013) note that information goods are hard to value, especially due to their feature as an experience good and the subjective nature of valuation. Still, Yousif (2015) argues that data can be considered as capital, since it is valuable, produced good, even though the value is vague. This sets an interesting basis for data monetization, when the value of data varies, it is hard to estimate and is fully realized only after the use.

Many authors argue, that data should be seen as one of the most valuable asset for companies (Yousif, 2015; Manyika et al., 2011). Data itself can be

valuable because of different factors and context: It can be used to optimize the logistics services and to personalize user experience on e-commerce sites (Rossman, 2016), to optimize the repair and service solutions for jet engines (Smith, 2013; Mackenzie et al., 2014) or even to encourage people to live healthier with data-driven insurances (Gore et al., 2017). Spijker (2014) presents the value of data and its competitive advantage with Apple's iTunes service: Previously music distributors lacked all the important meta data considering the music consumers, whereas with digital services such as iTunes, the music provider knows who bought the music, when and where, and can combine all this information into unique profiles of preferences. As can be observed from aforementioned examples, the value of data is case-specific: In some cases the data can be used to benchmark company's performance, understand the business environment better, differentiate commodities, decrease costs by optimizing value chain or by personalizing user experience (Spijker, 2014; Rossman, 2016). As Yousif (2015) mentions, for data-driven processes and products, data is the fuel that makes them run.

In this thesis, value of data relates to the *value derived from using data in different use-cases*. Value of data is relative to individual's perception and context, therefore it is case-specific. Due to data being an experience good, the value of data in a specific context can be realized fully after the use of it in the context.

2.2 Drivers for increasing data usage

This section presents the different drivers and enablers of increasing data usage. The changing business environment considering big data and supporting technology and legislative development are discussed as well.

Due to the trend of digitalization, the focus of companies is turning more into data (Spijker, 2014; Brown et al., 2011; Fred, 2017; Yousif, 2015; Manyika et al., 2011). Spijker (2014) argues, that the trend is moving especially from hardware to data ecosystems in multiple traditional industries. In these ecosystems, the products act more as a medium to create data, than as an end product (Spijker, 2014). This is supported with the discussion around *big data*: Big data and the possibilities beyond it, have been widely discussed and hyped in academic as well as business literature (Brown et al., 2011; Chen et al., 2012; Buhl et al., 2013; Davenport et al., 2012). Big data is mostly used as a term to describe potentially more insightful data (Fred, 2017). Big data is defined with three V letters: Volume, Variety and Velocity (Chen et al., 2012; McAfee and Brynjolfsson, 2012), but Value and Veracity are often added as well to describe the nature of data (Ylijoki and Porras,

2016; OECD, 2013). One evidence of the trend is the significant growth forecast for big data and analytics market: Investors forecast big data markets to grow from current $122 billion revenue of 2016 with a compound annual growth rate of almost 12% until year 2020 (IDC, 2016). The current transactions and customer interactions already create vast amounts of data, and the embedded sensors of Internet of Things, social media and data through devices such as mobile phones accelerate the growth of data (Lewis and McKone, 2016; Brown et al., 2011; Davenport et al., 2012). Therefore, companies posses significant amount of data, which changes the business environment in different ways: With big data, there is an increased transparency, more possibilities for collaboration, more incentives for experimenting, better customization of products and more evidence driven management (Brown et al., 2011; McAfee and Brynjolfsson, 2012; Manyika et al., 2011; OECD, 2013). For example, Brown et al. (2011, p. 1) describe: *"In fact, big data may ultimately be a key factor in how nations, not just companies, compete and prosper."* Thomas and Leiponen (2016, p. 15) discuss about big data ecosystems and notes that *"Data ecosystems will profoundly disrupt businesses in nearly every consumer-centric industry"*. Multiple other authors acknowledge the potentially broad impact of big data on multiple different industries (Chen et al., 2012; Mayer-Schönberger and Cukier, 2014).

The constantly growing amount of data creates more possibilities for data utilization, and the created data tends to create more data (Yousif, 2015; Opresnik and Taisch, 2015; Spijker, 2014). The new opportunities to utilize data are driven by multiple patters: Mackenzie et al. (2014) and Spijker (2014) mention the explosion in the amounts, but also on the content and quality of digital data: There is vast amounts of real-time connected data available due to drivers such as increased connectivity and mobility of users, Internet of Things and digitalization of processes. For example, location-based data is much more accurate and can be accessed instantly, which opens up new opportunities to leverage the data. The decreased cost of sensors combined with increasing amount of openly or commercially available data, ensure that the variety of data is not the bottleneck for utilization (Spijker, 2014). Therefore, the amount and quality of data enable further utilization of it. Najjar and Kettinger (2013); IDC (2016) argue that the change in business environment has been driven not only by the vast amount of data, but also by the supporting analytical capabilities and cloud computing. The cloud environment has enabled new virtual services and possibilities to distribute services and value to different actors much easier and faster. The better capabilities and basis to analyze, exploit and share data have made new innovative ways to monetize data possible (Mackenzie et al., 2014; Najjar and Kettinger, 2013). Due to popularity of digitalization and big data, data

monetization has emerged as a topic as well (Fred, 2017).

As one industry-specific driver for greater data utilization, the Payment services Directive 2 (PSD 2), Directive (EU) 2015/2366, changes the legislative environment of finance industry (European Comission, 2015). For companies operating in financial sector in Europe, the legislative changes due PSD2 will force the companies to revisit their strategy and position considering their data (Cortet et al., 2016; Salmony, 2014; Haataja, 2015). Due to PSD2, banks are required to provide application programming interfaces (API) to their data, which increases the possibilities and threats of other actors utilizing and monetizing company's data (Cortet et al., 2016; Salmony, 2014). This opens up new possibilities for financial technology, fintech, companies to collaborate with established banks, as they can create new services and innovations based on the access to the bank accounts (Salmony, 2014). Often these new actors can provide a wide variety of personalized features, which increase customer intimacy and decrease the attractiveness of traditional financial institutions (Spijker, 2014). This accelerates the transformation of current companies, as they need to decide their strategical position in the new business environment (Cortet et al., 2016; Haataja, 2015).

2.3 Data monetization

Data monetization as a phenomenon has emerged due to the aforementioned changes in business environment as well as understanding the value of possessed data. This section explains and defines the concept of data monetization, and what are the different options for companies to monetize their data. Therefore, the nascent offerings of data monetization are presented and clarified with examples of data monetization from different global companies. Finally, a summary of ways to monetize data is presented.

2.3.1 Definition of data monetization

The concept of sharing and trading data is not new (Moore, 2015; Mackenzie et al., 2014; IDC, 2016). Nevertheless, the term *data monetization* is currently fairly ambiguous: Monetization is about utilizing something, data in this case, as a source of profit or to coin something into money (Fred, 2017; Merriam-Webster, 2017). Najjar and Kettinger (2013, p. 213) describe data monetization as *"when the intangible value of data is converted into real value, usually by selling it"*. They still include conversion to other tangible benefits and to reduced costs as monetization as well. As Fred (2017) points out, different definitions use the output of monetization as profit or

revenue. She notes that data monetization lacks a specific definition and defines data monetization as *"the revenue generation with and out of data and data-derived and information-based products and services"* (Fred, 2017, p. 24). This definition includes data-derivations; Products and services created out of data are included in data monetization as well. This is the definition of data monetization that is used as a basis in this thesis. To narrow the scope further, the focus of this research is on the data companies currently possess, not the potential data that they could have in some point in the future.

Data monetization being a non-core business utilizing data created in core business has been presented as an idea in multiple articles (Fred, 2017; Najjar and Kettinger, 2013; Thomas and Leiponen, 2016; Woerner and Wixom, 2015). However, academic research and practitioners' focus have been mainly on the possibilities of big data within an organization (Thomas and Leiponen, 2016): Majority of big data and analytics discussion focuses on internal possibilities and dismiss the external possibilities, or mention monetization or commercialization of data only shortly. This can be seen in multiple popular articles, such as Brown et al. (2011) and Davenport (2015), where it is briefly acknowledged that data can be used to create new products and services in addition to improved internal decisions. Therefore, the data monetization as a term and phenomenon is not yet widely studied (Fred, 2017). Fred (2017) argues how some definitions define monetization as generation of money flow, which is interpreted as creating revenue or making profit. Data monetization can be considered as converting intangible value of data into real value (Najjar and Kettinger, 2013). This can be done by selling the data or the intangible value can be converted into other tangible benefits, such as discounts and advertisements. Thomas and Leiponen (2016) discuss about *data commercialization*, where data is monetized as an asset, rather than analyzed and used to improve existing or new products and services. Therefore, data monetization and commercialization relate to the same phenomenon, and data monetization is used as a term in this thesis.

Spijker (2014) discusses in his book about data monetization and data driven strategy: To monetize data, some other actor has to value it. As Thomas and Leiponen (2016) mention, the value of data monetization is rarely created in isolation, but more interdependently and mutually with other stakeholders of data ecosystem. Companies can leverage their data gathered in operational processes to create new value with others. This can be due co-creation with others by letting others to co-create with company's data (Spijker, 2014). Yet, the equivalent perception on the value is needed for both sides of monetization (Fred, 2017). The increasing availability of external data, especially open data, has enabled new possibilities for combi-

nation of internal and external big data. The term *data reuse* has been used to describe the secondary use of the gathered data, often from more a technical point of view (Thomas and Leiponen, 2016; Alkkiomäki, 2015; Opresnik and Taisch, 2015). Traditionally byproducts of operations have been commercialized in different ways (Lewis and McKone, 2016), and there's some research on exhaust data, where the focus is often on non-core data of organization (Thomas and Leiponen, 2016; Manyika et al., 2011). The initial purpose of exhaust data might relate to operational control, quality checks or reporting, but it might still possess some value for different actors after use (Opresnik and Taisch, 2015; Spijker, 2014). Data monetization as a term does not restrict the initial source of data, and therefore data reuse relates to data monetization as well. The following table, *Table 2.1*, presents the search results and popularity of different terms. Different terms were searched with search engine Google, academic search engine Google Scholar and academic literature database Scopus. The search terms used were defined, for example "big data" or "data monetization". The results of data monetization and data commercialization includes the results from terms "data monetisation" and "data commercialisation" as well, since the spelling depends on the used language group, UK or US. As can be observed in the results, the terms data monetization or data commercialiazation are not widely used. Data reuse as a term is more popular, but relates often to more technical perspective of data usage. Big data as an emerged trend has been widely studied and results a significant amount of search results.

Table 2.1: Popularity of different data monetization related terms. Search done with chosen search engines on 30.08.2017.

Used search term	**Search results in Google**	**Search results in Google Scholar**	**Search results in Scopus**
Big data	69 800 000	213 000	34 527
Data monetization	134 600	407	18
Data reuse	176 000	16 900	1 187
Data commercialization	16 250	236	3

Depending on the academic source, data monetization can focus on internal data utilization as well: For example Wixom and Ross (2017) divide data monetization into three different categories: 1) Improvement of internal business processes and decisions, 2) Enhanced core products and services with

information and 3) Information offerings sold to new and existing markets. Najjar and Kettinger (2013) as well include reducing the costs with data into data monetization as well. As Thomas and Leiponen (2016) note, majority of authors discuss on the internal processes and current products and only present briefly the idea of *new business models*, or *information offerings* in this case. As Fred (2017) notes, data monetization has been used as an umbrella term to define all actions which aim to generate revenue with or out of data or data-derived products or services. These examples demonstrate, how the definitions of data commercialization and data monetization haven't been generally defined, but the definitions and content depend on the specific use cases. However, Gartner evaluated in 2014, that by 2016, 30 percent of businesses have begun directly or indirectly monetizing their information assets (Huang and Laney, 2014).

This thesis focuses on the data monetization, where monetization creates value for other stakeholders as well, not just the initial organization. Therefore, in this thesis, an external actor is involved in data monetization: Data is monetized with different stakeholders, and the idea of improving internal processes, services or decision-making through data is not considered as data monetization in this thesis. In this thesis, the data monetization is based on the definition provided by Fred (2017, p. 24): *"The revenue generation with and out of data and data-derived and information-based products and services"*. However, this research aims to provide a comprehensive definition for the term. Therefore, the data monetization is considered in this thesis as *the creation of revenue based on the possessed data*. In this definition, there are always at least two actors involved: Actor monetizing owned data and the actor using the data. The focus of this thesis is on the external monetization, which is divided further into different business models in the following subsection.

2.3.2 Offerings of data monetization

To understand the offerings of data monetization, it is important to understand the concept of business model. The definition of business model as a concept lack a clear definition (Zott et al., 2011), but can be considered as "the content, structure and governance of transactions designed so as to create value through the exploitation of business opportunities" (Zott et al., 2011, p. 6). Osterwalder and Pigneur (2010) present The Business Model Canvas to describe the business model of a business in one page. All required aspects are gathered to one canvas, such as key partners, key activities, customer segments, channels and revenue streams. Still, on the center is the value propositions for customer, which can be seen as the core of business

model (Johnson et al., 2008; Osterwalder et al., 2014). This is true for data monetization as well: Data has to contain value for customers for them to pay for it. Since data monetization is a novel phenomenon, the business models of data monetization differ as well. Najjar and Kettinger (2013) note, that there are not yet identified best practices considering data monetization. Therefore, the focus is on customer value proposition and data monetization offerings companies can provide. The other parts of business models, such as key resources and key processes (Johnson et al., 2008), are not widely discussed in the theory due to novelty of phenomenon as well as scarcity of supporting literature.

One significant part of the business models, and closely linked to offerings, is the revenue model (Zott et al., 2011). Najjar and Kettinger (2013) acknowledge the need for suitable contracts, trust, data packaging, pricing models as well as marketing models. Thomas and Leiponen (2016) mention different options for pricing the data monetization solutions. One option is a *freemium* pricing, where free basic service encourages to use the service and the revenue is created by a premium pricing for more detailed data. Additionally, subscription fees for access to data, pay-per-use model for API usage and advertising can be used as pricing options (Thomas and Leiponen, 2016). Pricing and the revenue stream are vital parts of business models (Koutroumpis and Leiponen, 2013; Zott et al., 2011), but they are not analyzed further in this research. This is due the novelty of the researched phenomenon and different business models related to monetization, as well as the restricted focus of the study. Since the prevalence of data monetization is still vague, it is more important to focus on the strategical reasonings behind the monetization than the practicalities of packaging or pricing offerings.

In this section, different offerings for data monetization are presented. Companies can monetize their data by providing value to their current customers or to different actors, third parties. This is acknowledged in offerings, since the changing end-customer affects business models and the offerings in them as well. Therefore, the data used for monetization, can be gathered in one operational process and monetized by providing value to the same customer or another one.

Spijker (2014) notes that the successful use of data can create truly competitive advantages, for example Amazon's highly targeted recommendations and Mint.com's personalized expensive management, create differentiated customer experience when compared to traditional competitors. Still, the successful monetizations of data are rarer to find, as the data marketplaces and data brokerage markets are highly fragmented (Faria et al., 2016; Thomas and Leiponen, 2016). Wixom and Ross (2017) presented the concept of in-

formation offerings as a way to create value for customers from data, whereas Mackenzie et al. (2014) identified five different data derived categories: 1) Augmenting Products to Generate Data, 2) Digitizing Assets, 3) Combining Data Within and Across Industries, 4) Trading data and 5) Codifying a Distinctive Service capability. Yet, only Combining Data Within and Across Industries and Trading data can be considered as data monetization, since the others focus on how to turn current assets or processes into new products, or how to generate more potentially valuable data. Combining Data Within and Across Industries tackles the coordination of information across industries and sectors with data integration. By coordinating information and co-creating with city authorities, companies and customers can create new value to all involved stakeholders from data. Trading Data is about monetizing current data directly with other actors. By exchanging data with others, it is possible to create new collaboration platforms which create accurate products by combining data from different sources. Moore (2015); Huang and Laney (2014) divide customer data monetization into direct and indirect monetization: Data can be sold or traded directly, or data can be used to create new information products or services, where data is monetized indirectly. However, Fred (2017) approaches data monetization from three different layers of refinement: Unstructured data, structured data and information-based products or services. As can be observed in presented examples, the business models and offerings considering data monetization are vague and opaque. Yet, there seems to be a mutual understanding that companies evolve through different phases in their commercialization activities (Thomas and Leiponen, 2016; Najjar and Kettinger, 2013). Thomas and Leiponen (2016) note that companies progress from experimentation with data into more collaborative and complex business models. Offerings can be divided into different categories depending for instance on the purpose of data use or the format or maturity of monetization. Still, companies and their purposes can differ significantly. Therefore, the monetization is divided in this research to three different categories by the nature and maturity of offerings:

1. Selling data
2. Providing insights or analyses and
3. Creating a scalable service or a product

As Spijker (2014) and Thomas and Leiponen (2016) note, none of the patterns or data monetization offerings are superior to others, but depend on the chosen data strategy. Yet companies progress from simpler monetization experiments into more complex ongoing ones. Both of the authors acknowledge the reality, where companies create value by using more than

one approach simultaneously. This is noted in this research as well, since the presented categories are used to represent and clarify the phenomenon of data monetization. Therefore, these categories do not aim to cover all data monetization offerings, since every company form a unique and complex unity with different customers, products and company's legacy.

2.3.2.1 Selling data

The simplest way to monetize company's data is by selling it (Spijker, 2014; Thomas and Leiponen, 2016). Organizations can provide raw, less differentiated data for others to reuse as a data supplier (Thomas and Leiponen, 2016). Sometimes data need to be analyzed, repackaged or anonymized further (Spijker, 2014). Nevertheless, the data is sold, in a raw form or more aggregated one. By parsing, cleaning or cataloging data, company can act more as a data manager and increase the value through transformation of data (Thomas and Leiponen, 2016). Wixom and Ross (2017) argue that monetizing the data by selling is the hardest way to draw value out of it. This is mainly due to unique business model, which isn't directly linked to current core products and services, where the data is gathered. As Mackenzie et al. (2014) note, data monetization solution is rarely a sustaining innovation, but more often a disruptive one. Companies can sell their data to current customers or suppliers, but this kind of behavior is often integrated to current relationships and offerings, such as Vendor Managed Inventory (Angulo, 2004). Therefore, selling data can be seen as a part or additional feature on current services or it can act as a distinct service aimed for new customers.

If company sells its data, it generates the least amount of potential revenue of monetization, since the raw data is rarely refined much further (Thomas and Leiponen, 2016). In industries of fragmented information, raw data sharing is sufficient, and does not pose risk for the initial business model (Thomas and Leiponen, 2016). Data aggregators, who combine and analyze data from multiple sources, are one potential customer for raw data (Thomas and Leiponen, 2016). The attractiveness of data aggregators has increased as the data masses and sources have multiplied (Brown et al., 2011; Thomas and Leiponen, 2016). For example in health care industry, data aggregators can integrate clinical, behavioral, public-health and payment data to create more accurate insights for treatments (Brown et al., 2011). There has been little research on data marketplaces and trading of big data and insights derived from it to external actors (Thomas and Leiponen, 2016). Gartner assume that by 2020, 25% of large organizations will be either sellers or buyers of data through formal online data marketplaces (Faria et al., 2016). The data trading has been already a big business for illegal activity (Yousif, 2015; Holt

and Lampke, 2010), even though it is not popular for majority of companies.

Organizations can pick up the data generated by their core process and sell it as a product to their current customers or third parties. Here are presented few examples how companies can sell their data:

- Telecommunications operator Vodafone sells its anonymized network data to navigator company TomTom (Mackenzie et al., 2014). Vodafone has real-time, location-based, data about its customers, from which TomTom can understand better movement on roads. This helps to optimize navigation, when there's more accurate information about traffic congestions.

- A transport company sells investors its real-time data considering global product shipments. This way investors can involve data considering marine vessels' movement and harbors' cargo data to their current business and economic forecasts (Brown et al., 2011).

- PatientsLikeMe, a social media for patients, sells anonymized data for partners such as pharmaceutical companies and medical device manufacturers. Data is patient-created and consists of information and experiences related to illness and treatments. (PatientsLikeMe, 2016)

- Toyota sells traffic data generated from cars to municipalities, corporate delivery fleets and city authors. The data users can utilize the bought data in infrastructure development and route optimization (Lewis and McKone, 2016)

- Ebay sells anonymized transaction data generated in its platform to interested third parties. This way the data originated from customer activity can be exploited to create additional revenue (Opresnik and Taisch, 2015)

In these examples, data is valuable for the companies due to different factors. Vodafone's data monetization provides real-time location data for TomTom, which can improve its own products with the data. Since the traded data is provided as raw data, it can be integrated easier to used systems. For investors, the real-time knowledge of cargo movement and harbor activity can improve the accuracy of valuations of different geographic locations. In PatientsLikeMe case, pharmaceutical companies and medical device manufacturers get valuable information about the actual life and usage of their end-customers, patients. This is the reasoning in Toyota's case, since authorities get real usage data on road infrastructure. In eBay's case, third parties are provided with accurate and real-time information considering consumer buying behavior.

2.3.2.2 Providing insights or analyses

Companies rarely just sell their data directly to third parties, as Spijker (2014) notes. Instead companies can provide data-based insights or analysis (Thomas and Leiponen, 2016; Spijker, 2014). These insights can carry information considering customer insights, such as segments, habits, interests and plans, as well as advertisement targeting and payment analysis (Thomas and Leiponen, 2016). By providing customers data-derived analysis, company can add more value, for itself and its customers (Spijker, 2014). In addition, handing over original, or even anonymized data, might compromise company's business (Spijker, 2014). Thus, companies can provide insights and analysis derived from data, without allowing other companies to access the original data. The provided insights and analysis can be provided to limited chosen partners without creating a scalable service. By limiting the potential customers, companies can control the delivered value, and can avoid the issues considering identifiability of single data points or unknown reuse of data. There are some examples how companies have provided insights from their data:

- Pharmaceutical distributor Tamro provides drug manufacturers and other suppliers insights about customers' spending regarding their drugs in specific locations (Tamro, 2017b).
- Tamro also offers its customers, pharmacies, information regarding their sales, products and other variables compared to their competitors (Tamro, 2017a).
- Barclays Bank sells anonymized retail benchmarks to UK retail chains, when compared the spending pattern of banks' customers' in different stores. Bank has valuable data about its customers' consumption behavior, which can be leveraged for other actors (Spijker, 2014).

In all of these cases, the initial data is historical. Still, the insights derived from data provide valuable information about customer behaviour, that the buyers of data insights would otherwise miss. Therefore, companies can understand their competitors and the business environment better, which leads to more data driven decision-making. Information such as *share of wallet*, recency, frequency and monetary value of purchase when compared to competitors are valuable information when compared to competitors (Spijker, 2014). By providing insights, analysis or benchmark, there are less risks involved than by providing the original data itself. In addition, companies can limit the monetization offerings only to chosen partners.

2.3.2.3 Creating a scalable service or a product

Companies can also create new value-adding services and platforms to scale the delivery of data or insights (Thomas and Leiponen, 2016; Spijker, 2014; Najjar and Kettinger, 2013). This can be implemented with dashboards or similar interfaces (Whitmore, 2016). The more complex models often create more revenue, but are normally harder to execute (Thomas and Leiponen, 2016). Initially monetization solutions can be tested with few actors, but eventually scaling is needed to provide sufficient revenue (Najjar and Kettinger, 2013). Data and value derived from it can be provided to multiple customers through a scalable service, but the control over the data and what it is used for weakens. Spijker (2014) presents 'commodity swap', the transformation from commodity services into value-added services, as one data monetization option. In commodity swap, the sale or usage of a commodity product is used to generate data, which is then used for differentiation and to create new offerings from data. As an example, electricity doesn't differ in quality, but power networks can carry data about consumer behavior (Spijker, 2014). With smart meters, it's possible to identify the specific machine used in a household. This opens up new opportunities to monetize the collected data, when electric utility providers can provide appliance manufacturers accurate data about their appliances usage. Mackenzie et al. (2014) note the vast potential with smart meters: Gathered data can be sold to appliance manufacturers or to be used to create new enhanced services to homeowners. As Spijker (2014) points out, this could result even electricity utilities leasing machines for customers according to their usage. This is rather similar business model as Rolls-Royce's *Power by the Hour*, where customers are paying for the used flying hours, not the whole jet engines as products (Smith, 2013). There are some examples how companies have created a new innovative and scalable service from data:

- Google's smart thermostat product Nest monetizes the collected data by providing it to electric utilities (Dillet, 2014). Nest doesn't share actual data with utility providers, but provides them a chance to balance their energy grid. The provided insights about energy, appliance and utility usage are all packaged in the service, so the utility companies don't get hold on the data (Rossman, 2016).

- Foursquare has created Foursquare Analytics, where users' location data is analyzed and provided for other companies (Dent, 2017). Service utilizes data from Foursquare's location database, which is used by multiple different companies, such as Snapchat and Twitter. With

this data, Foursquare can provide businesses information about their customers' consumption behavior, demographics and even routes.

- Barclays Bank provides a service platform for SME companies, where companies can get insights about business inflow and outflows and more analyzed data about payments and transactions (Barclays, 2016). This data is compared to similar businesses in similar locations, so SMEs can benchmark their figures to averages.

- Adara, a software company, uses data from multiple sources to provide advanced customer segments and profiles to partners from partners' data. This way partners, such as airlines, hotels, and travel agents, get refined information, which can be used to provide highly-specific services and advertisements (Spijker, 2014).

In all of these examples, the value for data owner comes from multiple customers. Online portals or dashboards can ensure, that companies get personalized insights, but the original data itself doesn't transfer. By creating a platform, such as Foursquare Analytics, companies can increase the value of their service by adding more stakeholders to the service as well as more data sources. This way companies can leverage the capabilities of different actors, as the platform owner does not need to have all the required analytical capabilities when there are other actors involved (Najjar and Kettinger, 2013). This kind of platform can create a data flywheel effect, where the increased data creates even more data (Yousif, 2015; Spijker, 2014; Rossman, 2016). Adara connects different data providers and refines their data more accurately as more partners start using the service. Thus, the service benefits from the increasing amount of users and data sources, as the accuracy and variety of data improves as well.

2.3.3 Summary of data monetization options

As described earlier, the data monetization offerings and business models are still mostly vague. There are different stages of data monetization, where the maturity of monetization solution varies from selling raw data to interacting with multiple actors through a data-driven platform (Najjar and Kettinger, 2013; Thomas and Leiponen, 2016). In most cases, interacting in the value chain or value network is the most natural way to monetize data (Spijker, 2014). This is similar to forward or backward integration (Amit and Zott, 2012): Data flowing through a value chain or network can create more value, when it is combined with other data from different parts of the chain. Organizations in the busy spots of data traffic can try to benefit economically from

the access to the data flow (Thomas and Leiponen, 2016). In some cases, companies might not be in the same value chain, but they can still benefit from sharing data. This is often due to the shared final customer (Spijker, 2014). Sharing information about customer behavior and combining it with own data can enhance company's understanding of its customers.

The following table, *Table 2.2* summarizes different options to monetize data according to the theoretical review.

Table 2.2: Different ways to monetize data according to the literature. Summary based on the theoretical review.

Customer / Offerings	**Provided to current customers**	**Provided to other actors**
Selling data	Data is sold for current customer. In supply chains, companies can provide their suppliers raw or aggregated data about inventory levels of suppliers' product.	Data is sold for a third party actor. Vodafone can provide navigation companies such as TomTom anonymized data from its mobile network to improve TomTom's real-time navigation system.
Providing insights or analyses	Insights are provided to current customers. Pharmaceutical distributor Tamro can provide its customers insights considering their sales compared to competitors.	Insights are provided for a chosen third party. Cargo companies can sell insights to investors about commodities based on marine vessels' movement and harbors' cargo data.
Creating a scalable service or a product	A more comprehensive and scalable data-driven solution created for multiple customers. Barclays Bank provide SME companies a service, where they can analyze and compare their financial KPIs to other similar businesses in similar locations.	A service or platform created to connect different data consumers. Google provides insights of Nest smart home usage to utility firms through a service, where they get energy, appliance and utility aspects of home users.

2.4 Issues affecting data monetization

Data monetization is linked to company's core business, industry and data itself. Since data monetization is closely linked to utilization of data, some of the issues derive from the utilization of big data. Since data has some fairly unique features, such as it can be copied perfectly, it is often combined with other sources, it can be used simultaneously by multiple users and is difficult to protect, there are different issues considering the monetization of data (Thomas and Leiponen, 2016; Yousif, 2015). This section presents the most significant issues affecting the data monetization and utilization of big companies according the literature. The section is divided to different subsections depending the issue: *Organization related issues*, *Industry related issues*, *Data related issue* and *Privacy, legal and security aspects*.

2.4.1 Organization related issues

The biggest issues considering big data utilization are not technology related (Bean, 2017; Spijker, 2014). Spijker (2014) notes that most people in organizations often recognize the value of data and consider leveraging it. Nevertheless, bigger organizations are designed to be stable and consistent product-driven businesses. Since data monetization can be about innovation in new markets for different customer segments, the company's rigidness prevents data monetization. Concerns related to the data commercialization are rarely technical, but instead cultural (Thomas and Leiponen, 2016; Bean, 2017; Lewis and McKone, 2016). Haug et al. (2013) notified in his study, that the biggest barriers considering master data quality, are due to organizational issues. These includes barriers, such as lack of sufficient roles and responsibilities, inefficient organizational procedures, lack of management's focus and support and missing procedures and quality measurements. Bean (2017) evaluates that the biggest barriers in data utilization are related to cultural challenges, such as organizational alignment, lack of understanding data and change management. Often companies already struggle sharing information inside the organization, as business functions support organization wide poorly (Laney et al., 2015). Data monetization requires new processes, skills and resources, and a suitable culture to sufficiently support the creation of new offering (Wixom and Ross, 2017; Spijker, 2014). Brown et al. (2011) underline the need for sufficient talent: The demand for people with deep analytical skills in big data as well as managers and analysts is much higher than there are suitable people available in the United States. Najjar and Kettinger (2013) present different pathways to data monetization, regarding

company's technical and analytical capability: Depending on the situation, companies can build capabilities, hire or exploit a third party, acquire more data or start monetizing with partners.

Clear strategy is needed, as well as sufficient investment and commitment from organization (Wixom and Ross, 2017; Laney et al., 2015). Accountability for monetization, and lack of it, delays and prevents the monetization of data (Wixom and Ross, 2017). Without committed managers, the redirection of employees for new value creation doesn't happen. Therefore, data-related new projects should be regarded as projects of strategic importance, since they transform company into more data-driven organization (Spijker, 2014). The need for data driven strategy or data strategy is noted by multiple authors (Spijker, 2014; Wixom and Ross, 2017). A successful monetization as an innovation requires a strong business-unit leader with a sufficient team, since the monetization is often about creating and growing a new line of business (Wixom and Ross, 2017). This is supported with studies considering innovation in big companies: A dedicated team, autonomy, use of informal networks, room for experimentations and risk-taking are needed instead of rigid structure and legacy culture (Govindarajan, 2016; Wessel, 2012; Stringer, 2000; McDermott and O'Connor, 2002). If data monetization solutions are analyzed through the traditional business channels, they can be often regarded as high risk and low profit (Spijker, 2014). This combined with the non-core nature of data monetization makes it more difficult to establish monetization solutions (Fred, 2017). As data monetization can transform the current business model, it takes determination and courage to create monetization solutions (Spijker, 2014; Najjar and Kettinger, 2013). It isn't sufficient to provide a right data set and tools for employees, but the company culture, organization structure, employee capabilities and processes and habits need to support the chosen data monetization business model (Wixom and Ross, 2017).

2.4.2 Industry related issues

The industry, where a company operates, affects the possibilities of data monetization. Naturally there are different options as well as issues affecting data utilization, when company gathers data considering consumers, or when the data is related to industrial equipments. OECD (2013) address the differences between different industries: Some might be data-intensive, yet are under-exploiting their data, whereas some are less data-intensive and have smaller growth potential regarding data. As OECD (2013) note, some specific industries are more data-driven than others: Financial services, communication and media, utilities, government and discrete manufacturing are

especially data intensive. Brown et al. (2011) studied the value potentiality as well as ease of capture considering value creation for different industries in the US. Finance and insurance in addition to information sector and government had the biggest potentials for benefiting from big data. Still, for the government sector, the possibility to capture value from data was significantly low. Other potential gainers from big data were wholesale trade, real estate and health care providers. The construction, manufacturing and administrative services had lowest potential for value creation from big data.

The industries affect the potentiality of data utilization as well as the ease and premise. For example, it is harder to monetize data with partners, if the partners are not mature enough nor their systems support modern data analysis. Yet, the evaluation of potential gains as well as current capabilities differ between different authors: Brown et al. (2011) place manufacturing in the bottom for potentiality, whereas in big data and IoT researches, the manufacturing industry has been seen as an industry with huge potentiality (Visconti et al., 2017; Opresnik and Taisch, 2015). The B2B companies often possess different data, mostly on their products, whereas consumer-facing companies might have valuable information about consumers. For example, a telecommunications firm can sell their data generated from customers' movement (Thomas and Leiponen, 2016). And as Yousif (2015) mentions, for financial service companies, their data asset is a gold mine. However, finance service companies face stricter legislative restrictions than manufacturing companies regarding the usage of data (Evans et al., 2012). Thus, the industry affects the potential use cases of monetization, as well as the barriers.

2.4.3 Data related issues

Data quality is one of the first problems companies face when starting to monetize data (Spijker, 2014; Wixom and Ross, 2017). Data quality can be measured by multiple variables: Fisher and Kingma (2001) mention that accuracy, timeliness, completeness, consistency and data relevance are the most used variables. Accuracy means that the gathered data reflects the real facts or value and lacks errors. Timeliness means that the recorded data should not be out-of-date. Naturally in some cases the data can be older, but in more demanding cases, there have to be recent data available. Completeness implies that all variables are recorded and retained, and not one is missing. Consistency refers to the consistent representation of data and lack of redundancy. Relevance refers to the applicability of data for the particular issue, so that the data can be directly applied to the problem. Spijker (2014) mentions the same variables, except has uniqueness as an external variable. This means that there are no duplicate values in the data,

but this idea is included in consistency in earlier definition. In addition, he discusses about the importance of data reliability. Since data can originate for example from social media behavior, it is important to address the reliability of data if it is monetized.

Accessibility and quality of data are one of the most significant obstacles to monetizing data (Wixom and Ross, 2017). The lack of access and the insufficient quality can prevent the monetization greatly. According to Wixom and Ross (2017), only a quarter of companies offer employees and customers easy access to the needed data. Data liquidity refers to getting the wanted data in the required format with minimal time, cost and risk (Yousif, 2015). This can be a significant issue for companies, especially in data monetization, when data derives from one business function or operation and is used in another. As Wixom and Ross (2017) mention, you cannot monetize data that no one can access. In many cases, the purpose of generated data have not been to be sold or used outside the organization (Spijker, 2014). Therefore, data is often created in operations processes, the accuracy and features of data are not optimal for the final product. When the data itself is the product, the requirements for the quality of data might be higher when compared to internal use of data. Therefore, the data quality need to be covered sufficiently (Thomas and Leiponen, 2016). The combination of trivial issues form a complex environment, which results in a need to establish structured and continuous process of data management and supply (Spijker, 2014). Fitness for use summarizes the data quality: High-quality data is fit for use by data consumers (Strong et al., 1997). As Fisher and Kingma (2001) mention, data is then presented in a format which serves the user's purpose and need, and data is stated in terms familiar to that user. In addition, data quality is relative and case-specific: Some data might be useful for strategic planners, but it might not be suitable for operational engineers.

Data continuity is another issue to consider with data (Spijker, 2014). The majority of data monetization products rely on continuous delivery of data for customers. Like Spijker (2014) defines, data continuity is the process that ensures that the raw material, in this case data, is consistently transformed into a product. Yousif (2015) raises the issue of data equality, which relates to capturing and keeping data in its original shape and format. This is good to acknowledge as well, when more and more data is packed and analyzed, so that the content isn't distorted nor any valuable aspects lost. Variations in production, different campaigns, seasonality of business and changes in used IT systems interfere the equality and continuity of data flow (Spijker, 2014). The used technologies and tools need to support this and seamlessly integrate with each others (Yousif, 2015).

2.4.4 Privacy, legal and security aspects

When data includes personal customer data, privacy concerns are extremely important (Thomas and Leiponen, 2016; Moore, 2015). In consumer-driven business, data is often customer-centric, which naturally increases privacy concerns (Spijker, 2014). Still, Thomas and Leiponen (2016) note that privacy concerns are present in most of the industries, if not all. When more and more data is traded, the role of privacy increases as well (Najjar and Kettinger, 2013). Privacy issues can prevent data monetization, for example for mobile operators, since they are afraid of compromising customer experience or losing their trust (Thomas and Leiponen, 2016). As Spijker (2014) notes, perfectly legal and anonymized data products still receive criticism from the public: ING Bank presented legal service, for targeted ads based on transaction data. Only few days later it faced criticism and pressure from consumer organizations, banks and even the national parliament (Spijker, 2014). Irobot, company behind Roomba robot vacuums, experienced similar weakening of public brand: Company's CEO mentioned the possibility to share and monetize customer data, collected by vacuums in people's homes, with smart home applications (Wolfe, 2017). This was followed by a public discussion, after which the CEO was quick to rephrase and correct the idea (Kastrenakes, 2017). Therefore, even though some monetization service is legal to establish, it does not mean that consumers or media will appreciate it.

Ethics around data, especially personal one, are important to consider (Huang and Laney, 2014; Fred, 2017). As Ohm (2010) notes, computer scientist can often *reidentify* or *deanonymize* individuals in anonymized data sets. Even though there is transparency in privacy agreements, the lengthy and complex policies overwhelms individuals more than inform them (Thomas and Leiponen, 2016). Nevertheless, consumers seem to value the benefits of the additional services compared to the decreased privacy (Spijker, 2014; Moore, 2015). Tension between convenience and privacy is evident and should be acknowledged properly (Brown et al., 2011). These issues are broad, and have different weight in different industries: If company collects data about individuals' movement or about the weather, the privacy might have a different role.

The legal environment affects the potential monetization. There are different industry-specific legislations, that prevent direct monetization of data. Many authors, for example Moore (2015), Weber (2010) and Thomas and Leiponen (2016) address the need for regulative changes to support the changing data environment and technologies. Yet, companies need to be careful regarding privacy and changing regulations (Huang and Laney, 2014). Some

legislative changes, like GDPR, affect the utilization of company's data.

The General Data Protection Regulation (GDPR) approved in European Union in 2016 and to be implemented in 2018, does create tension on data monetization. The goal of this Regulation 2016/679 is to harmonize data privacy laws across Europe as well as protect and unify EU citizens' data privacy (European Parliament and of the Council, 2016). GDPR gives users a *rights to be forgotten* and request erasure of personal data as well as right to retrieve their data from a service and share it to another (Thomas and Leiponen, 2016; Mantelero, 2013). For companies these regulations cause caution: Sanctions for companies not complying to the directive can be as high as EUR 20 million or 4 percent of annual revenues (Blackmer, 2015) as well as any information related to individuals are considered as personal information and should be handled as such (Rich, 2016). Currently regulatory complexity prevents the creation of data ecosystems, since there are limited mechanisms of legal frameworks, policies and guidelines for coordination (Thomas and Leiponen, 2016). The GDPR is controversial, as it changes the requirements, definitions and there are not yet any precedents, which causes companies to approach data utilization more carefully (Rich, 2016).

In addition to legislative changes, the legislation depends on the geographical area, which affects possibilities and requirements for utilization of data (Thomas and Leiponen, 2016; Fred, 2017; Huang and Laney, 2014): In European Union, there's a transposition of the directives in each of the 28 member states, whereas in the United States there are 50 different states with potentially different laws regarding IT security and privacy protection issues (Thomas and Leiponen, 2016). In addition, the value of data might be restricted to geographical areas, as Finnish industry information might not be considered valuable abroad (Fred, 2017). Additionally, the regulations are dispersed in different industries as well (Huang and Laney, 2014), as can be seen in PSD2.

Data security is a key requirement for big data in addition to aforementioned legislation (Thomas and Leiponen, 2016; Fred, 2017). It's important to ensure that applicable policies, regulations and laws are complied, while value is still created (Yousif, 2015; Chen et al., 2012). While more data is easily available due to rising amount of new devices and sensors and the data flows into cloud storages, the concern for data security rises as well (Brown et al., 2011). As Weber (2010) notes, there are privacy and security threats created by IoT solutions, which need to be addressed properly. The increased amount of integrations of IT architectures create new opportunities for external parties, while risks rise related to security and intellectual property (Brown et al., 2011). Security and data governance are tightly linked

together: Authorization, access, auditing and encryption are issues to be considered of (Yousif, 2015). IP protection of data varies greatly in different countries (Thomas and Leiponen, 2016). In addition, the legal instruments might protect the structure and organization of the database, not the data it contains (Thomas and Leiponen, 2016). Thomas and Leiponen (2016) note that data agreements considering IP protection, regulatory complexity and pricing concerns need to be resolved for data monetization. These relate to trust, which is also an important factor to consider when external parties are involved in monetization (Najjar and Kettinger, 2013). Trust and the ownership of data can raise issues and potential conflicts between different stakeholders, as companies often aim to own the data and control the value creation (Thomas and Leiponen, 2016; Fred, 2017; Moore, 2015; OECD, 2013; Bilbao-Osorio et al., 2014).

2.5 Summary of literature review

The aim of this literature review was to summarize and present relevant academic and business literature considering data monetization. Three different research questions were presented in *Chapter 1.* This chapter covered all research questions in this thesis, yet the the aim was to specifically answer research question RQ1, and establish a basis for empirical research for research questions RQ2 and RQ3. The focus of this chapter was to explain and define data and data monetization for this thesis, as well as present how companies can monetize their data and which factors affect the monetization according to studied literature.

RQ1: What does data monetization entail for firms today?

To define data monetization, it is important to understand the concept of data and nature of it. In this thesis, data is considered unrefined set of discrete objective facts. Data has different features, such as being non-rivalrous, non-fungible and an experience good. Information and knowledge are seen as refinements of data. Value of data is subject to use case and context: Data being an experience good affects the valuation. Nevertheless, to monetize data, it has to hold value for other actors as well. Data monetization still lacks a specific definition, and in this thesis data monetization is considered the creation of revenue based on the possessed data. In this definition, there are always at least two actors involved, company monetizing data and another company using it. The current changes in business environment support the monetization and utilization of data: The big data, increased connectivity, mobility of users and IoT ensure the sufficient amount of high-quality data,

whereas cloud and analytics capabilities enable the refinement and sharing of data. In finance sector, the PSD2 encourages companies to utilize their data broader.

RQ2: How do firms monetize their data?

The data monetization is often created mutually with other stakeholders; the current customers and partners being the most obvious ones, but companies can monetize their data by providing value to third party actors as well. Therefore, the monetized data can originate from one operational process and be provided to the same customer or another ones. This affects the novel business models of data monetization. Yet, the offerings can be divided to selling data, providing insights or analyses and creating a scalable service or a product. Companies can provide initially simpler monetization solutions, such as trading data or insights for limited partners and evolve their offerings into more complex services provided to multiple actors. These complex services are often more scalable and can benefit from multiple users using the service and creating even more data. Still, the offerings of data monetization are hard to separate, since the data monetization markets are nascent and the solutions can be integrated to current services and customer relationships.

RQ3: What affects data monetization done by firms?

Like in data utilization, data monetization is affected by a lot of different issues as well. The monetizing organization and its culture create a basis for new service creation, and the risk aversive mentality, the lack of focus and support from management with rigid organization can prevent the creation of monetization solutions. Naturally there are industry specific issues, which make some industries more potential and capable for monetization than others. However, the first issues companies face, relate to quality and accessibility of data. Data need to be high quality enough, easily and cost-effectively accessible, fit for the use as well as a continuous data flow and creation is required. Therefore, the data infrastructure need to be in a good order for potential monetization. In addition to organizational and data issues, the legislative restrictions and changes can prevent monetization of data. Privacy is a sensitive topic, and need to be approached carefully. In addition, the changing legislative environment with GDPR transform the possibilities for data monetization: GDPR causes new caution for companies with significant possible penalties, yet legislative changes can accelerate the use and sharing of data for other actors, as it is with PSD2.

Chapter 3

Methodology

This chapter describes the research methodology used in this thesis. The research setting, process and used methodology are presented. Additionally, the data collection and data analysis processes are explained regarding the empirical parts.

3.1 Research setting and process

This study explores the phenomenon of data monetization in different Finnish companies. This study and topic were initiated by digital service company Futurice, for which this thesis was performed. The concept of data monetization had emerged increasingly in company's customer base, since customers were looking for new ways to utilize their data. This set the basis for the study: The aim was to map the field of data monetization and understand and describe how big companies are currently using, sharing and selling their data. As Saunders et al. (2009) describe, the approach to the research problem can be exploratory, exploratory as well as descriptive, or descriptive. In this research, the approach is exploratory as well as descriptive, since the nascent phenomenon is explored, but still the aim is to describe data monetization and the surrounding business environment.

The research process started with a broad literature review to understand the concept of data monetization. It was evident, that the term was not clearly defined nor widely studied, as the similar practices were studied under terms of *data commercialization* or *innovative big data practices*. Yet, the initial literature review provided sufficient basis to narrow the research problem and formulate research questions and limitations to guide the research. As Birkinshaw et al. (2011) mention, in qualitative research a careful approach can be taken, as the topic should be evaluated before choosing what

to study. After defining the research problem, the research plan, process, required tasks, deadlines and schedule were formulated. *Figure 3.1* presents the initial timeline and planning created for this thesis. During the research process, plan and schedule were refined and validated to ensure a successful research project.

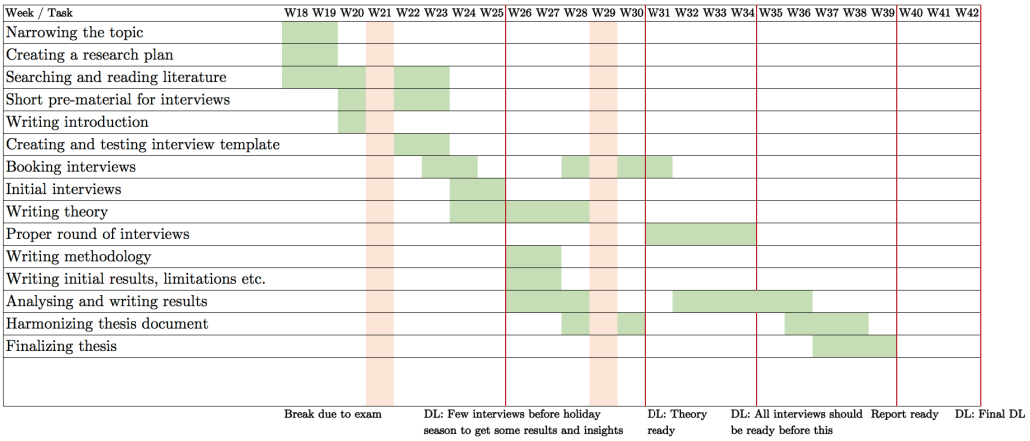


Figure 3.1: Initial plan for research timeline.

After the research plan was formulated and process decided, a more narrow literature review was conducted focusing on data monetization, especially data monetization as a concept, its business models and issues preventing it. This provided a satisfactory basis to understand the concept and business environment before empirical research. Then it was possible to formulate pre-reading material for interviews, interview structure, book interviews and eventually hold interviews with different companies. Eventually results were analyzed, reflected on theory and summarized.

3.2 Research methodology

This section describes in detail the methodology used in this research. First the qualitative research method and the research approach are presented, as well as the inductive approach. Next the chosen research strategy, case study, as well as the chosen research method, interviews, are presented.

3.2.1 Qualitative research method

Since the focus of the study is to explore and describe a novel phenomenon, a qualitative research approach was chosen for this thesis (Saunders et al., 2009;

Ritchie and Lewis, 2014). As Ritchie and Lewis (2014) describe, qualitative research can be used to describe the nature of what exists, examine the reasons and associations between what exists, appraise the effectiveness of what exists as well as aid the development of theories, strategies and actions. Birkinshaw et al. (2011) note that qualitative methods are exploratory in nature and meant to be used to approach weakly understood phenomenon. Therefore, the research questions presented in *Chapter 1* can be answered and data monetization described more broadly with qualitative research methods. As Birkinshaw et al. (2011) mention, the qualitative research allows a more in-depth understanding of studied phenomena. In addition, Robson (2007) notes, that the exploratory study is suitable to understand the phenomenon better and assess it from new perspective. Data monetization as a practice is not entirely new, since data brokers have sold data, but especially due to changes in business environment described in *Section 2.1*, it is a potential strategical choice for companies operating in other industries as well. Yet, it has been weakly studied in current context of modern business environment (Thomas and Leiponen, 2016; Fred, 2017). Therefore, this research combines exploratory as well as descriptive approaches, which can present accurate profiles of situations or phenomena (Robson, 2007). The used approaches support the purpose of the study, while as Saunders et al. (2009); Ritchie and Lewis (2014) note, the exploratory, explanatory and descriptive research can still be well planned and performed, even though the studied subject is fairly unknown. Nevertheless, Ritchie and Lewis (2014) mention the need for flexible research design in qualitative research, where iterative research methods can be used.

Theory on data monetization is very scarce (Fred, 2017; Thomas and Leiponen, 2016; Najjar and Kettinger, 2013), so inductive reasoning was used in this thesis. Inductive approach starts with the collection of data and then explores the results to see which themes or issues come up (Saunders et al., 2009). Therefore, evidence is used as the genesis of a conclusion (Ritchie and Lewis, 2014). Yet, theory was used to guide the qualitative research and create a basis for interviews in this research, but still the approach is not deductive, where the research is based on theories and frameworks and evidence is used in support of a conclusion (Saunders et al., 2009; Ritchie and Lewis, 2014). Inductive reasoning is a suitable approach for exploratory research, and can be used when there is no clearly defined theoretical framework to which reflect on (Saunders et al., 2009), which is the case in academic and business literature of data monetization.

3.2.2 Case study

A case study is a suitable method to induct theory in qualitative research (Eisenhardt, 1989). A case study is a strategy for performing research, where multiple sources of evidences are used to conduct an empirical investigation of a chosen phenomenon (Robson, 2007). As Eisenhardt (1989, p. 534) describes, *"case study is a research strategy which focuses on understanding the dynamics present within single settings"*. Case study strategy is a good tool to gain a rich understanding of the phenomenon as well as the context (Yin, 2009). Case study is especially suitable strategy for explanatory and exploratory research (Saunders et al., 2009), and therefore used in this research. In addition, case study is suitable for inductive theory building (Birkinshaw et al., 2011; Eisenhardt, 1989).

In this research, the multiple case study approach was chosen. This was due the novelty of the concept of data monetization, as well as the suitability of multiple case study approach to exploratory research (Yin, 2009). Yet, as the topic is novel and the practices and issues affecting monetization vary in different industries, multiple industries and companies as cases were chosen to ensure a broad exploration and understanding of the subject. Since the approach of this research was exploratory, the goal was to cover a broad scale of different Finnish companies to ensure a sufficient understanding of data monetization. Thus, a single case study would have described the phenomenon in a selected case, not explored the field of monetization. Still, the focus was not to draw generalizable results, but to explore nascent phenomenon in Finnish business industry. Therefore, the non-probability sampling, where samples are based on subjective judgment (Saunders et al., 2009), was used in this research. The non-probability sampling, such as self-selection sampling is a suitable sampling method for exploratory researches according to Saunders et al. (2009) and Ritchie and Lewis (2014). Self-selection sampling as well as heterogeneous sampling were used to identify central themes and a possible variation between studied companies (Ritchie and Lewis, 2014). Since the population used was identified as fairly heterogeneous, a relatively high sample size of cases is needed (Saunders et al., 2009). Yet, the sample consisted of multiple companies from similar industries, and the focus was on the big companies, established in Finland, which increased the homogeneity of sample. In addition, the logistical and resource limitations affected the potential size of sample, which was acknowledged in the initial research plan. Therefore, the sample size was decided to be from 15 to 20, eventually resulting in 19 different companies.

3.2.3 Interviews

Case studies often collect qualitative evidence from interviews and observations (Eisenhardt, 1989). Interviews are commonly used in qualitative research (Ritchie and Lewis, 2014), and were used in this research as well. Interviews provide a possibility to explore and describe phenomenon (Saunders et al., 2009). As the term data monetization is still vague and lacks clear definition, the interviews were a suitable method to collect reliable data on data monetization practices. For example with a survey, the reliability and validity of data might have compromised, even though the sample size would probably have been higher and wider. By interviewing 19 different companies, it was possible to explore the field of data monetization broadly: The evidences cover multiple successful examples on data monetization as well as multiple industries and companies, which have not monetized data. Since there were already identified differences between researched companies, the sufficient sample size ensured that there will be enough data to draw insights considering data monetization in companies and industries. This research maps the current situation of data monetization for Finnish industries, it can be considered cross-sectional study (Saunders et al., 2009). Yet, since interviewed companies are at different maturity of data monetization, the results can be utilized at later stages as well, when the snapshot on the phenomenon is broad enough.

Interviews were performed as semi-structured interviews, where the themes and questions were covered, but varied on the interviewees and companies. Saunders et al. (2009) mention that semi-structured interviews are suitable to exploratory as well as explanatory researches. In interviews, the organizational as well as business context were acknowledged: For example, with companies operating in finance industry, the PSD2 regulation was discussed about, since it is tightly linked to data monetization strategies. In addition, companies with current monetization solutions were asked to describe the solutions, whereas companies without monetization practices were asked more about the issues preventing the monetization. The interview questions were divided to four different categories: Introduction, Current data and data utilization, Data monetization and barriers and Wrap-up. Open-ended questions were used to ensure the high variety and information richness of answers. The used interview structure can be found as *Appendix A*.

As Saunders et al. (2009) propose, interview themes were created for the semi-structured interviews. The interview themes and questions were formed according the information acquired from literature review as well as from a few preceding interviews with industry experts. These interviews were used to understand the current academic field of data utilization and

monetization before the data collection. These interviews were conducted as unstructured interviews, which are informal and non-directive, where the interviewee is given the opportunity to talk freely about the topic (Saunders et al., 2009). For data collection and interviews, a short pre-reading material was created. This material was meant to define data monetization and gave a list of examples of data monetization. This way interviews very easier to promote, while the material defined the novel term and framed the discussion for interviews. This material was sent to interviewees before the interviews, but it was not used to guide the interviews. The pre-reading material slide-set can be found as an *Appendix B*. A brief research on companies, their data utilization and digital business was conducted before the interviews. This consisted of reading company web pages, news articles and company blog posts considering data usage and digital solutions. All interviews were conducted in Finnish, except one. Due to the exploratory approach of the study as well as the novelty of phenomenon, all the interviews were conducted face to face to ensure reliable findings. Interviews consisted of the interviewee and interviewer, except in one interview there were two interviewees present. All interviews were recorded, and due the sensitivity of the subject, the companies and interviewees were anonymized.

3.3 Data collection

The qualitative data was collected by holding 19 interviews with 19 different companies, one interview per company. First a long list of potential companies were formed. As mentioned in the earlier section, the non-probability sampling was used to choose different companies. This selection was guided by multiple factors: The business where company operates, the data that they potentially possess, as well as the potential value and uniqueness of data. For example, retail and media companies possess data on consumer behavior, whereas construction companies can collect data on buildings and infrastructure. As mentioned in *Chapter 2*, finance and insurance, information, government, wholesale trade, real estate and health care providers are potentially the biggest winners from big data utilization, according to (Brown et al., 2011). The final selection of potential companies was driven by this information, author's experience, discussions with people involved in data driven projects, as well as company websites and news considering data utilization, data strategy and data monetization. Due to the exploratory approach of the study, specific focus industries were not chosen. Nevertheless, the aim was to interview multiple companies from the same industry to draw some insights from the potential state of monetization as well as the

frequency or the gravity of faced issues for data utilization. The focus of the study was on big companies, for example publicly listed companies. Still, it was acknowledged, that the size by revenue varies in different industry and cannot be directly compared to each others, such as for media and manufacturing companies. Finally, the sample contained 44 different potential companies to interview.

Out of the identified 44 companies, 35 were contacted through email and phone. The networks on Futurice, author's own personal networks, company web pages and social media site LinkedIn were used to find the suitable person from chosen companies. Since this study approached data monetization from more strategical approach than technical, the people in roles of Chief Digital Officer (CDO), Chief Information Officer (CIO) or directors regarding data utilization were approached. If there was information on a more suitable person for the study, for example a person mentioned in news article regarding data utilization strategy, that person was approached instead. Out of the contacted 35 companies, 19 different interviews with 19 different companies were arranged. The interviewed people had following titles, presented in *Table 3.1*. People working as directors had titles such as *Director of New Ventures, Head of Digital Solutions, COO or Chief Data Scientist*. The interviewees regarded as other contained for example an interviewee working as a *Senior Data Analyst*.

Table 3.1: Interviewed persons, by title.

Title category	**Amount of interviewees**
Director	9
CDO	4
CIO	4
Other	2
Total	**19**

The interviewed companies are presented in *Table 3.2* and *Table 3.3*. Due to the sensitivity and strategical approach of interviews, the anonymity of interviewed companies is preserved. Therefore, only the industry and revenues of companies are presented.

Table 3.2: Interviewed companies, by industry.

Industry	**Amount of companies**
Media	4
Manufacturing	3
Finance	3
Health care	2
Forest	1
Industrial services	1
Telecommunications	1
Retail	1
Construction	1
Electricity	1
Logistics	1
Total	**19**

As can be observed from *Table 3.2* and *Table 3.3*, there are different sized companies from different industries present. This was acknowledged as a part of the exploratory nature of the research. Still, this affects the generalizability, validity and reliability of findings. These aspects are addressed in *Discussion*.

Table 3.3: Interviewed companies, by size.

Revenue in 2016 (Million €)	**Amount of companies**
>1500	7
500 - 1500	4
250 - 500	6
<250	2
Total	**19**

3.4 Data analysis

As Gioia et al. (2013) note, in inductive research, the interviews and analysis are often aligned. Therefore, the analysis started already after the first interview. First after an interview, the data was transformed from written notes into digital ones in separate text files. The audio recordings were listened through to ensure that no information was lost in the process and the points were understood correctly. After writing the notes, single interviews were summarized and then categorized into emerged themes. The themes used

in interview structure were initially used, but were refined as the amount of data increased. This kind of iterative analysis is usual for exploratory research (Saunders et al., 2009). In summarizing and categorizing results, coding was used to analyze different interviews. In coding, different labels and categories were created, and data was summarized and color coded from different interviews using single letters under the corresponding theme. *Figure 3.2* shows an example of used coding in this research. The used methods are similar to grounded theory methods (Saunders et al., 2009), since open and selective coding were used while summarizing and categorizing data.

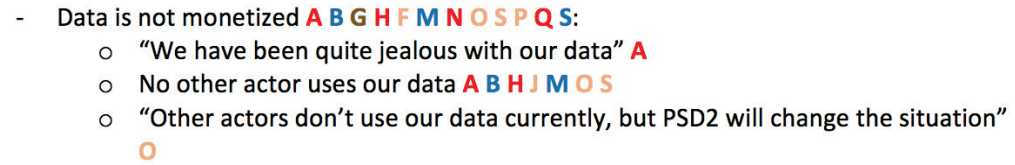


Figure 3.2: Example on data coding.

Data display and analysis approach can be used in inductive analysis as well (Saunders et al., 2009). Data was first summarized and simplified, after which it was organized and assembled into visual displays, graphs and tables. This is called data displaying, where matrices and networks are used (Miles and Huberman, 1994). As Saunders et al. (2009) note, data displays help to recognize relationships and patters as well as draw conclusions from data. Therefore, to understand the data monetization of different companies operating in different industries, different visualizations were used. For example, the maturity of company's data utilization was analyzed based on the observations on the interviews as well as the interviewees opinions considering company's ability and current usage of data. The criteria was formed and company's data utilization reflected on these criteria. From this analysis, a following graph, *Figure 3.3* was created. The initial interviews were checked to validate the position of a company on the different graphs related to other companies and their positions. The companies close to each other, were categorized into different groups. In this case, there were three identified groups: Companies with a low level of data utilization, companies with medium level of data utilization, and the companies with high level of data utilization. The criteria used for this example can be found in *Section 4.3*.

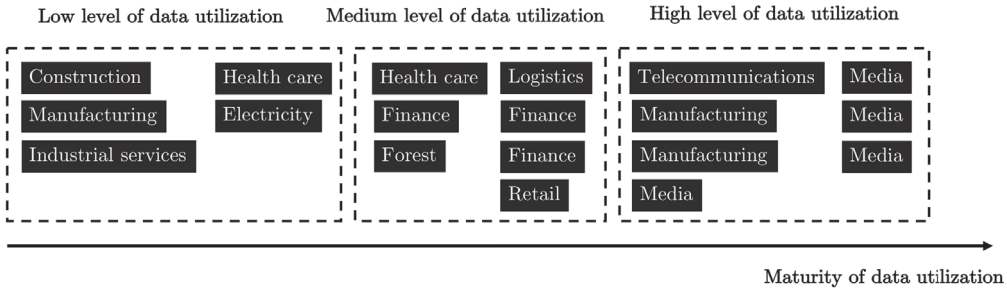


Figure 3.3: Example on data visualization.

Of different ways to analyze qualitative data, presented by Miles and Huberman (1994), noting patterns and themes, clustering, counting, making contrasts and comparisons, factoring and making conceptual coherence were used. Different quotes were used to represent different themes and patterns, but to make contrast and comparison as well between different data observations. Qualitative data was also quantified, for example by counting the frequencies of data monetization practices. Quantified qualitative data can be very useful supplement to the principal analysis of data (Saunders et al., 2009).

Chapter 4

Results

This chapter presents the results from the empirical research. The results are divided to four different sections: First the value of data and changing business environment are discussed to provide basis for other themes. Then the current utilization and monetization of data in interviewed companies is discussed. At last, the barriers preventing the monetization are discussed and summarized.

4.1 Value of data is hard to evaluate

The value of data is often relative to subjective evaluation and depends on the use case of data. As majority of companies have not monetized their data broadly, they are struggling to evaluate its value. This section discusses how the value of data is linked to industry and the use case, how companies value their data and how it is realized after the use of it.

The industry, where company operates in affects the potential value of data. The information considering consumers buying habits, financial capabilities, movement and change in life situation can potentially be utilized in many different ways: For example personalized information considering individuals can be combined with other data regarding the same individuals. This kind of approach can result in accurate consumer profiles, which can be exploited for example in marketing and sales to target customers as well as in R&D to understand the customers better. For companies operating more in a physical environment, the value of data consisted of information regarding the surroundings, geographic location and environment. Therefore, the value of data is closely linked to industry the company is operating in: Manufacturing, construction and industrial services companies had data considering environment and physical surroundings. As companies move from physical

environment into digital one, the value of data increases as well. Companies in finance, telecommunications, media, logistics or retail industries possessed valuable information about consumer behavior. The consumer data was regarded more valuable than other data, as there are more evident use cases for accurate data considering consumers.

The value of data is strongly linked to its use case and context. This was addressed by multiple companies, as one interviewee working in a media company noted: *"The value of data is directly linked to what it is used for"*. The raw data was perceived rather low in value, but with refining or combining data into other data increased its value. For example manufacturing companies perceived their data lower than other interviewed, as the direct applicability of data to other contexts was not as evident as it was for companies possessing data about consumers. The increased value by refining data was addressed by multiple companies in different industries: Raw data was rarely seen the source of value, but algorithms and analysis created the real value. Often the data was combined to other data sources to create more valuable data. Since all companies collected data used for ordinary business issues, such as transaction data, information about customers and financial data, raw data was seen valuable mostly when it was unique to the business environment: If there are not other actors possessing and providing similar information considering environment or consumers, the value is often higher, when there are less competition for providers of such data. Therefore, the valuation of data depends greatly on the use case of it.

The perceived valuation of company's data was evaluated with following criteria:

- Perceived internal value of data: How did the interviewee perceived the value of company's data? How valuable is the data for company?
- Perceived external value of data: How did the interviewee perceived the value of company's data for other actors? Is the data evaluated to be valuable for other companies as well?

Following figure, *Figure 4.1*, presents the results from interviews regarding the perceived value of data. Companies were grouped into three different categories, depending on how valuable they perceived their data. The first group only perceived the data valuable for the company, the second group estimated the data to be valuable for other companies and actors as well. The third and the biggest group evaluated the data to be extremely valuable for companies, while being valuable to others as well. The industry related value factors are present, as the more digital companies regarded their data more valuable.

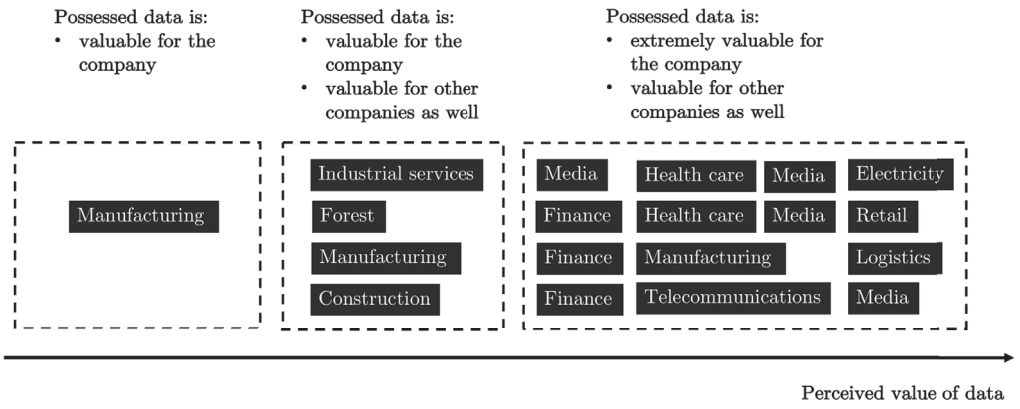


Figure 4.1: The perceived value of data.

The value of data is realized after the use of it, and without monetization practices, it is hard to evaluate it in those use cases. Almost all companies estimated their data valuable for other companies as well, while few companies operating with consumers saw the collected data as their most valuable asset. This high valuation of company's data is natural, since no company wants to diminish the potential value of their data and business. The high valuation creates a suitable premise for monetization, as it is hard to monetize data, that is not considered valuable. Yet these valuations are perceived value by the interviewed company, since the real value of data is only realized after the use of it. Therefore, the value of company's data for other companies is subjective evaluation for most of companies. As majority of companies do not have validated use cases of external usage of their data, the valuation of data and the potential benefits of monetization are rough estimations.

4.2 Changing business environment accelerates new data usage

The changing business environment accelerates the broader use of current data. The increasing amount of data, supporting cloud, analytics and data science capabilities, pressure from markets, PSD2 and GDPR all accelerate the broader usage of company's data. This changing environment results in better possibilities for companies to monetize their data.

Significant amount of data

As multiple authors, such as McAfee and Brynjolfsson (2012) and Chen et al. (2012) note, the amount of data gathered by established companies is

nowadays huge. This can be seen in the results as well: None of the interviewees felt that they did not have enough data. Only a few company out of 19 interviewed needed some data that they did not possess yet, which was due to lack of technological solutions to gather accurate data. One manufacturing company described the amount of data by expressing that *"everything that could be imagined as data source, is used to gather data".* The data collection was rarely restricted and mostly all possible data was gathered. One interviewed CDO underlined this reasoning: *"The huge amount of data does not matter that much, since algorithms take care of the data mass. The problem is, if the data is not collected in the first place and there is not any data".* The supporting development of IoT accelerates the data amounts further, as sensors can be implemented to collect more data about physical environment and individuals' behavior. Therefore, the amount of data was rarely the bottleneck for data usage. As one interviewee working as a Head of Analytics described the situation: *"The problem is not the amount of data, but that the data has not been utilized properly."*

Master data, analytics and cloud capabilities

As Mackenzie et al. (2014) and McAfee and Brynjolfsson (2012) note, the business environment with increasing ways to capture, analyze, store and transfer data, increase the possibilities to utilize data. Yet, the opportunities are further supported by new capabilities, like one director acknowledged: *"Previously it was not possible to really utilize data. Now it is eventually possible and we can start focusing on how to use the data, and not just speculate at what we would hypothetically do at some point if we had the data or the right tools".* The proper data infrastructure is a premise for data monetization: Without suitable infrastructure, it is not possible to retrieve right data in the right form, nor build any data-driven services. For many companies, the proper data infrastructure is one key enabler for new possibilities around monetization: It has required a lot of effort to consolidate data from different business functions and sources to one place. This master data management enables a stable premise for data utilization.

In addition to master data management, the imrpved analytics, cloud solutions and technical capabilities accelerate broader utilization of data: The new technology supports the efficient analysis of data, when the software tools as well as algorithms enable the efficient and automatic handling of data while making the data richer in quality. The rise of data science teams with increasing amount of business-driven and technical analytics have enabled a better capability to provide data analysis services inside as well as outside the organization. This is further supported by the development around AI, which can decrease the amount of manual work required for data analysis

while providing more accurate results quicker. These contribute to decreasing costs related to data handling and analysis, as the data investments provide cost savings in the long run. In addition to analytics and AI, the development around cloud based solutions and APIs enable an easier way to share data, internally and externally. Proper APIs was a strategic initiative for few companies, as the benefits of APIs can be realized already between different business functions. Still, this creates new possibilities to use of aggregated data, also outside the organization. The improved analytics and cloud capabilities enable new ways to utilize collected data.

Pressure from markets

The pressure from markets and competition is a significant driver for data monetization solutions, as data monetization requires sufficient investments and focus from management. The investments were often due to pressure from markets: *"The impact of digitalizing environment is significant. It is a threat as well as an enabler for us to evolve"* described one interviewed director. The threat of disruption and digital transformation force companies to transform their businesses. This was further reasserted with comments from interviewees such as *we cannot afford to not do these data driven experiments"*,*"we cannot keep up with the competition, if we cannot build new services based on the data* or *"if we do not start utilizing our data better, this company will not be here in 10 years"*. To create suitable data infrastructure, with master data management and supporting tools, the support and investments from management is needed. The growing level of ambitions and goals considering data utilization supports the broader use of data. All interviewed companies stated directly, that the digital investments and the wider utilization of data are part of their current strategies. These were often brought up with evidences such as the role of CDO, a significant amount of new recruitments considering data utilization or by naming the digital topics as company's must-win battles. The pressure from markets drives the broader utilization of data.

Changing legislation of PSD2 and GDPR

The pressure from markets is further accelerated by the changing legislation, through PSD2 and GDPR. For companies operating in finance sector, the new Payment services directive (PSD2) transforms the business environment regarding data radically. Due to PSD2, companies are required to provide APIs to other financial actors. As companies pointed out, this encourages fintech companies, established and new ones, to create novel services based on financial company's data. One financial director underlined the change: *"Currently no other company uses our data, but the PSD2 will*

change this issue. Previously direct monetization was not possible, but PSD2 will encourage to data sharing and even to monetization." Therefore, the PSD2 increases the opportunities considering data monetization and encourages companies to reconsider their strategies considering data.

For companies operating with consumers, the General Data Protection Regulation (GDPR) by the EU Parliament sets new rules for data management and protection. Due to legislative pressure, companies are required to clarify their master data management, as they need to be able to provide individuals all the data company possess considering that specific individual. This requires investments from companies into data management and infrastructure, which will result in a better premise for potential new data-based services. These ways legislation can act as a driver for increasing data monetization.

Table 4.1 summarizes the criticality and frequency of enablers for companies discussed in this section. The enablers are evaluated based on the findings from interviews. Criticality refers to the gravity of enabler: How significant is the impact of the issue. Frequency relates to the popularity of enabler for interviewed companies. Both criticality and frequency are evaluated with the following grades: Low, medium and high.

Table 4.1: Enablers of monetization.

Issue	Description	Criticality of enabler	Frequency amongst companies
Significant amount of data	There is an increasing amount of data in companies, which is accelerated by IoT. This enables new ways to use the data, as the amount is not a restriction.	Medium	High
Master data, analytics and cloud capabilities	The improved data infrastructure, with proper master data management, new analytical capabilities and possibilities to utilize cloud environment, creates a good premise for monetization.	High	Medium
Pressure from markets	The pressure from markets drives new investments and focus of company's management towards broader data utilization. The threat of disruption acts as an incentive to utilize current data assets better.	High	Low
PSD2	PSD2 forces companies to rethink their data strategies, as established competition and emerging fintech companies can utilize company's data assets.	High	Low
GDPR	GDPR indirectly accelerates broader data usage, as it creates a better premise for data utilization. This way companies can utilize improved data infrastructure with new data-based solutions.	Low	High

4.3 Data monetization varies greatly, yet is not advanced

Data monetization varied greatly amongst interviewed companies. The data monetization solutions were commonly novel for interviewed companies. Even though the majority of companies had solutions, where their data was utilized by their customers or other stakeholders, these solutions were mostly operative cooperation to improve current products and services with current customers. Over a third of the researched companies had data monetization offerings, yet not many comprehensive ones. Majority of these data monetization solutions were yet nascent businesses and some were small scale experimentations, but a few companies had more mature services by which data was commercialized as an established open service. The following table, *Table 4.2* summarizes the popularity of data sharing and monetization amongst the interviewed companies.

Table 4.2: Data sharing and monetization in interviewed companies.

Practice of data utilization	**Amount of companies (N=19)**
Company's data used somehow by another company	14
Data monetization offerings provided to other companies	8

Over two thirds of companies let some other company to use their data to some extent. Yet, direct sharing of company's data, for example to partners, was less likely. As one media company noted: *"Companies have traditionally been quite jealous to share their own data, even though they use external sources of data quite a lot"*. This jealousy or reluctance to share company's data was brought up in all industries, as companies avoided sharing data due to potentially lost asset. However, sharing data can be used to create new strategical partnerships and to back up current ones. In about half of the companies, data was shared to some extent to other stakeholders, for example as aggregated data. Activities, such as sharing information to other parts of value chain, was used to enhance the operations. For example, data was offered to suppliers and partners for them to learn more about their products' performance, campaigns or sales. Though, it was pointed out, that when the value of data is high, one must be careful how to utilize and share it. The behavior was further underlined by one director: *"Our data usage is quite*

traditional: We buy data from other actors and use our own data, but we don't let others to use it". Therefore, companies let cautiously others to use their data, which sets a challenging premise for data monetization.

Following subsections discuss the current usage of data and data monetization practices in interviewed companies. It is essential to understand the current level of data utilization in companies as it sets a premise and context for data monetization.

4.3.1 Currently data is used to improve internal processes, products and services

The current level of data usage varied between interviewed companies. Data was used broadly in all companies for example to enhance internal processes, tasks and decision-making, and to improve current products, services and marketing. Yet, the level of data usage differed: Some companies were already very data-driven, whereas some had just started to use broader the possessed data. The level of this *maturity of data utilization* affects the opportunities and attractiveness of data monetization as well as the problems companies were facing. This maturity relates to the maturity of company collecting the data, as it is vital for company to be ready and mature enough regarding data utilization to create innovative data solutions, such as data monetization solutions. The maturity of data utilization helps to understand the current level of company's data usage.

The maturity of data utilization was evaluated with following criteria:

- The data capabilities: How did the interviewee perceived the company's capabilities for data utilization? Do they have skillful employees to utilize the possessed data?
- The estimated state of data infrastructure: In what kind of condition the data infrastructure is? Is it easy to access the data and how high is the quality of data?
- The data activity: How did the interviewee estimated the company's current actions related to data? What kind of data derived solutions or projects company had internally or provided as products or services?
- The ambition level: What kind of goals and aspirations company has set regarding data utilization?
- The use of external data: Does the company utilize external data sources?

The maturity was analyzed based on these criteria, yet driven by the observations on the interviews as well as the interviewees' evaluation. By

using these criteria, companies were grouped into three categories: Low level of data utilization, medium level of data utilization and the high level of data utilization. Low level companies did not have broad utilization of data and were just starting to utilize the current data in new ways. Medium level companies had already a broader level of data utilization, as they had better capabilities and had utilized data internally and created data-based services as well. The companies with high level of monetization were very data-driven, as data was used broadly and the data infrastructure with data science capabilities was mature. The evaluated maturity of data utilization in different interviewed companies is presented in *Figure 4.2*.

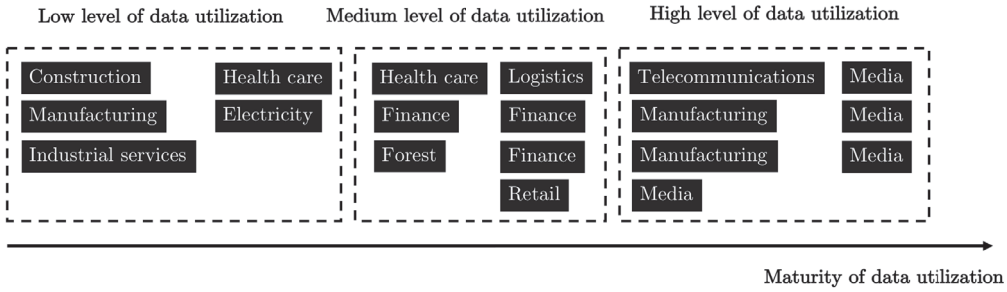


Figure 4.2: Estimated maturity of data utilization in interviewed companies.

Companies with low level of data monetization focused their data usage on internal processes: Companies needed first to create and optimize their data infrastructure, before it was possible to create new offerings derived from data. Therefore, the companies of low maturity were often focusing on internal data processes, and less on new services and products. As one manufacturing company's director noted: *"We are at the early stage of data utilization as our focus is still on internal projects."* The amount and level of current data actions were low in industries such as construction, industrial services as well as health care. In these industries, the data was not broadly utilized, as the data sharing and communication between different business functions and projects were low, practices were still traditional and very individual driven and the quality of data was relatively low. As Spijker (2014) noted, the data infrastructure need to be in a good shape before companies can build new value based on the data. Therefore, there were often still multiple internal data projects before company's data infrastructure was mature enough for broader utilization of data, for example with new data derived products. For manufacturing companies, data was used to optimize current processes to provide high-quality products with as low costs as possible. Yet, the proper utilization of data for new products was often missing in

different manufacturing and industrial service companies. The focus of data utilization was often in internal processes and decision-making, especially for companies which were beginning to utilize their data more broadly.

More mature companies focused their data usage on current offerings. With the supporting data infrastructure, companies could focus more on current customers instead of internal data usage. In consumer driven businesses, the products are increasingly commoditized as all companies provide seemingly similar offerings. This was the most evident in industries such as retail, telecommunications and finance companies, where the goal of broader data usage was often a possibility for differentiation. Thus, data utilization focused on the current products and services: The data provides new opportunities to personalize and differentiate the current offerings and provide a stronger relationship with the customer. For example in financing, customer receives similar value and service, regardless which finance provider's online service and portal is used. Though, the additional data-driven services and features, which can improve customer's perceived experience, can lead to a stronger customer relationship, and eventually to competitive advantage. Therefore, companies intended to use the data to lock in customers by providing a unique customer experience and value while personalizing the offerings. The focus of data usage of more mature companies was often on current customers through improved current services and products.

Majority of interviewed companies did not have monetization solutions, as their focus of data utilization was more on internal processes and current products and services. Yet, companies with suitable data capabilities and supporting infrastructure used data to create new services and products. Media companies were the most pioneering this, as they were the most mature in their data usage amongst interviewed companies. One media company illustrated this point clearly: *"Data enables directly or indirectly over half of our annual revenue. For example sales is fully data-driven as well as advertising."* This is natural, as media industry revolves around data, and the companies rarely possess heavy assets, such as physical business sites, factories or stocks of products. However, some manufacturing companies, which have aforementioned assets, had managed to create data-driven business. In these cases data was utilized broadly in multiple business functions, and was supported by high-quality capabilities such as distinct data science teams. The possibility to enhance traditional products with digital features had created new business and ever increasing amounts of data, which had created pressure to establish better data capabilities. By using external data, companies could enhance their own business as their offerings: Majority of interviewed companies utilized external data sources, such as registers and information providers. This was often related to high maturity of data usage,

as companies can utilize external data easier if it can be integrated to company's data infrastructure. If the data infrastructure and capabilities to use data are low, it is unlikely for companies to buy external data, if they struggle already to utilize their current data. Media companies were the most eager to buy external data, which was often combined and integrated to their own data. The focus of the most mature companies regarding data utilization was on the current products and services, as well as new ones derived from data.

4.3.2 Monetization solutions are often integrated into current relationships

The varying basis of data utilization guided the potential data monetization, as some companies still focused more on internal usage of data as well as improving current products and services. Yet, 8 companies out of 19 offered data monetization solutions, often provided to current customers on top of existing offerings. However, data monetization solutions were in some cases created as comprehensive new services to new customers, such as authorities, customers' customers and other third party actors.

Data monetization amongst interviewed companies is discussed in the following subsections. The subsections are divided based on data monetization offerings presented in *Chapter 2*: First the selling of data is discussed, then the provision of insights or analysis and lastly the creation of a scalable service or a product. After these subsections, the data monetization solutions are further discussed and summarized. *Table 4.3* presents the frequency of different monetization models between companies monetizing data.

Table 4.3: Data monetization in interviewed companies.

Practice of data utilization	**Amount of companies practicing monetization (N=8)**
Data monetized by selling data	5
Data monetized by providing insights or analysis	5
Data monetized by a scalable service	4

4.3.2.1 Data selling

Selling data as it is, or aggregated and anonymized, was performed by five interviewed companies, mostly media companies. As was discussed in *Chap-*

ter 2, the simplest way to monetize data is to sell it (Thomas and Leiponen, 2016; Spijker, 2014). Selling of data was realized in different ways: Some companies provided aggregated data through a service as a subscription basis, some as a single analysis or as an anonymized dataset and in sometimes data was provided as a raw data stream. Few media companies had APIs for third parties, the use of which was monetized. One company had a pilot project with a research company, to which anonymized, but detailed data on online activity of consumers was sold. Despite the popularity of selling data and possibilities to anonymize and aggregate data, interviewees emphasized the avoidance of selling the raw data. As Thomas and Leiponen (2016) mentioned, by selling data a company generates the least amount of potential revenue from monetization when data is not refined further. This was evident for many companies, as one CDO described the situation: *"Our data is not sold directly, but it comes with the product, for example through advertising"*. The reasoning behind this was often the strategical protection of valuable asset as well as less legal risks and the better price associated in providing refined insights and not the initial data. Even though five companies provided data monetization solutions as raw feeds, some of these solutions were not offered openly, but provided to partners only. This way the lost of valuable asset was less probable, as companies chose the partners strategically. This kind of approach was seen with one health care company, which did not share their data directly, but provided drug manufacturing companies anonymized and aggregated data on the use of their drugs. In offerings, where the raw data was openly available, data was often aggregated and anonymized to suitable datasets. By aggregating data, companies can act more as a data manager, and provide new value to current customers and partners.

The use case and customer of data guided the used medium: Customers of data monetization used the acquired data to improve their own business or products, acquire business intelligence, understand their competitors as well as to sense market trends. In some cases, the reasoning for providing aggregated or raw data, was due to integration of data into customers' systems. This was familiar for monetizing companies, since majority of companies monetizing data, had also bought external data. One manufacturing company provided raw data feed from their products for customers in different forms, depending on the customer preferences: The same data was provided as a raw data stream, a specific dataset or ready analysis through a service or an application, depending on preferences and use case. One potential customer for raw data is data aggregators and research firms, who combine and analyze data from multiple sources. Since the further combination and analysis, the data need to be as original as possible. One interviewed me-

dia company had a strategical partnership, where other company received anonymized data for integration to their own data. This way the analysis of data can be "moved" to other company, if the monetizing company lacks the capabilities to do the analysis. Though, company misses the potential added value created by refining the data. In most cases, the selling of data was provided as an additional service offered with other service. This way customer can analyze the data further if needed, but the use of raw or aggregated data is monetized. When data is integrated to used systems or further analyzed and explored by the customer, the raw data is the most convenient offering for customers. Yet, for company monetizing the data, it poses potentially lost business opportunities and risks related to ownership, legality and reuse. Despite majority of companies monetizing data sold it, companies avoided to provide data as it is or aggregated as a distinct service.

4.3.2.2 Providing insights or analysis

Companies aimed to refine the possessed data further, before providing it to customers. As Thomas and Leiponen (2016) and Spijker (2014) mentioned, by just selling data, companies miss potential revenue as data is not refined further into analysis. Due to this, mainly all companies providing monetization offerings, had solutions where customers get ready insights or analysis, but not the original data. The insights, analysis and reports contain the information valuable to customers, but do not cause issues regarding ownership or legislative restrictions of data or potential reuse of data for unknown use cases, as it is often in selling data. Companies preferred dashboards, portals or similar interfaces, where customers could get the information easily. These services could provide the needed information, such as insights considering real estates, consumer profiles, segments and their behavior, market demand, company informations, performance amongst decided segment or information regarding environment and surroundings. When providing insights or analysis, initial data could be combined to other data the monetizing company possesses or acquires. Media companies emphasized the role of combining data from different sources as a prerequisite for their data monetization. This way, companies can act more as data aggregators, who provide value by combining and analyzing multiple data sources. By enriching data with other data before analysis, companies can increase the value created and monetary value of offering. Still, enriching data was not common for other companies than media companies, even though the desire was evident.

The insights were often provided to current customers and partners, not as a distinct service for third party actors. For example, one media company gathered data from their customers, and provided insights back to the cus-

tomers considering their business environment by analyzing aggregated data. The aggregated data was provided for free for the customers to differentiate the service and provide personalized offerings. This way monetization offerings could be used as an incentive for new customers to commit more to the company providing monetization solutions, as the insights were offered as an additional feature and not as a distinct service openly to anybody. When insights are integrated into current offerings, they can be easier offered to suitable customers. By limiting the data utilizers to current customers, companies can control the delivered value as well as ensure higher revenue by bundling the feature into other offerings. As a further bundling, companies can provide consulting service with the provided insights, when the data is not traded as it is but as insights. By providing analysis service, companies can provide a more personalized service, while understanding better what kind of value customers gain from monetization offerings as well as other offerings company provides. Yet, only one company had established a data-based consulting service as an additional feature on top of their current data monetization solution. As the monetization offerings were integrated to current customer relationships, they could not attract new customers from different parts of value network or even from other industries.

4.3.2.3 Creating a scalable service or a product

Companies aimed to create scalable services to provide monetization solutions, yet these were rare. With an offered service or a product, companies can attract more users than with limited insights. In these cases, the service is often openly available as a distinct service, which can be used by any actor, not just current customers or chosen limited partners. Majority of companies providing monetization services aimed to establish their data monetization solutions as scalable services. This goal was addressed as well by Najjar and Kettinger (2013), who noted that scaling is needed to provide sufficient revenue. By providing information to multiple customers through a scalable service with portals or dashboards online, the customers get the valuable output from data but not the initial data. Yet, the service can include aggregated data as well: One interviewed company had a pilot project, where the scalable service acted as a platform, which connects different users of data to one service. The third party actors got data through APIs and the service encourages them to work on the company's data and provide new services and value for end-users. This kind of approach was not common.

Companies can also establish a service, where customers' data is enriched with company's own data. This was common for media companies, as they provide consumer profiles and segments, reflected on customers' data. This

was provided by other companies as well, which could reflect customers' data on own database, enrich it and provide it back to customer. This way customers of monetization can get their own data updated with external data sources, while the monetizing company does not need to provide all the original data. As the most advanced data monetization solution, one company had established a fully data-based service, where customers varied from data brokers to media and marketing customers. This specific data monetization solution was used to help customers to optimize their current marketing as well as to understand better their business and their customers. As in other services, the value was provided through dashboards or the customers' data was enriched on the service. When the value is provided in a more refined form as a service, the original data is not transferred while the customer receives the value of data.

4.3.3 Monetization solutions first on top of current products and services

Majority of companies monetizing data, enhanced current services and products. The monetization offerings were mainly created on top of existing customer relationships. The prioritization of company's data usage was often driven by first internal processes, then current products and services, and finally new service and products. As Spijker (2014) and Thomas and Leiponen (2016) argued, the data monetization offerings depend on the chosen strategy. This was evident for interviewed companies, since almost all companies monetizing their data provided value through multiple approaches. Yet, the data selling and providing of insights and analysis were often first experimented with limited amount of customers or partners, before creating a more scalable service derived from data. This is natural, as the value of monetization solution can be verified easier and with less costs by providing the data or insights, without creating a distinct service. Simultaneously the data can strengthen current relationships and provide a differentiated service for customers. Thus, the data monetization solutions focused often on current products and services.

To understand better the level of monetization in different interviewed companies, the maturity of data monetization was evaluated. The following criteria was used to evaluate the maturity of monetization:

- Actions of monetization: What kind of data monetization solutions companies had?
- Sharing of data: Did the company share data with other companies?

The maturity of data monetization in interviewed companies is mapped in the *Figure 4.3* below. Four different groups were identified, where two latter are companies monetizing data: Company not monetizing data, company not monetizing data, yet data is used externally, company experimenting on monetization solutions or providing small scale monetization and company providing comprehensive monetization solutions.

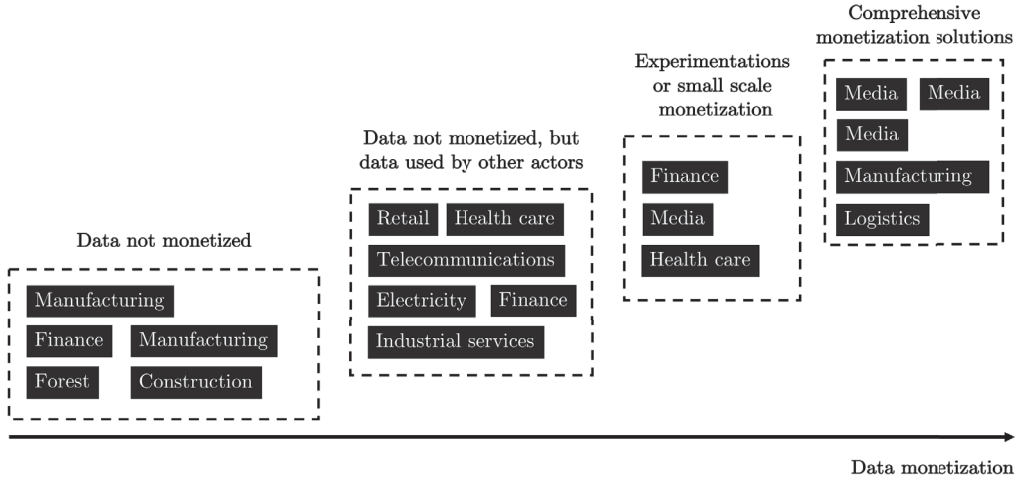


Figure 4.3: Maturity of data monetization.

The most pioneering companies in data monetization provided comprehensive monetization offerings. These solutions were often provided by media companies: As one media company stated, *"this is not just the early-stage blind enthusiasm, but we have already maturity and the aim of data monetization is clear."* The most mature companies considering monetization had clear growth ambition for monetization, while some had set sales targets and calculated repayment periods. Therefore, the state of monetization was already mature and was not considered as experimenting anymore, when there were pressure to create increasing revenue. The value of data monetization solutions provided by media companies were often due to combining and enriching data to provide accurate information for customers. In addition to data driven media companies, a logistics and a manufacturing company had established scalable monetization services. These were driven due to relatively unique high-quality data, which was valuable to multiple parties, and thus monetized through openly available online services.

Companies often first provided monetization solutions to current customers and partners before wide scalable solution. Companies with pilot projects, small scale monetization and experimentations on data monetiza-

tion implemented them with chosen partners. In these cases, the monetization solutions were often integrated to current relationships with customers, and were not available as distinct services. These customers of early monetization were often current customers of different businesses or found in current value chain. Therefore, the customers of monetization moved more from current customers to new customers, as companies matured in their data monetization. The development of early data monetization solutions was sometimes even accelerated by the customers' contribution: Some of the data derived services were customized solutions for specific customers and their needs. The benefit was validated value for customers, but the potential scalability of monetization solutions for other actors was unknown. Therefore, companies preferred to first create monetization solutions for their current customers, and not other actors. The importance of scalable monetization was still underlined: Especially solutions, which were co-created with customers and highly personalized, were aimed to transform into more modular and scalable solutions to provide more revenue.

Even though, some companies had pioneering data monetization solutions, the comprehensive monetization has not broken through. There are some companies, that are practicing it in a few ways, yet this is not common for Finnish big companies. Some of the companies monetizing data saw themselves as external refiners and sellers of data, whereas some saw monetization as a small external part of their business. For majority of companies, data monetization was not common business. One manufacturing company even emphasized the minimal potentiality of data monetization for their organization: *"I don't think that our data is valuable to any other stakeholder.* The expectations considering data monetization were careful amongst all companies: Monetization was noted to be small when compared to companies' core businesses. As one company with a novel data monetization service remarked: *"The direct monetization of customer data is a very niche market for us"*. Even for companies who had data monetization, the size of monetization markets was unclear: All the companies monetizing data had doubts on the size of markets for their monetization solutions. Even companies with few years on monetization experience were still unclear about the real size of business opportunities considering monetization. However, the potentiality was realized to some extent, as one director noted: *"Data monetization will not have its own division, but there are still easily millions of euros to collect with our data"*. Therefore, companies approached monetization carefully, often by focusing on current customers through improving current products and services.

4.4 Barriers preventing monetization

As presented in *Section 2.4*, companies have a lot of different issues affecting the utilization and monetization of data. Since the level of data monetization varied greatly between the interviewed companies, the gravity of some barriers and the extent to which they prevented monetization differed. For some companies with low maturity of data utilization, the barriers were still preventing the broader utilization of data, whereas for companies with ready data monetization solutions, barriers prevented the wide-spread scaling of their monetization offerings. The issues affecting the data monetization are divided to five different categories presented in this section:

- Potential benefits do not outweigh acknowledged risks
- Other data related projects are prioritized before data monetization
- Organization's culture nor capabilities are not suitable for monetization
- Data is not easily available nor in good shape
- Privacy and legal issues cause careful approach

4.4.1 Potential benefits do not outweigh acknowledged risks

The evaluated benefits of data monetization can be lower than the acknowledged risks. Spijker (2014) noted, that if data monetization solutions are analyzed through the traditional business channels, they can be often regarded as high risk and low profit initiatives. In this subsection, the high risks through position of trust, small evaluated benefits as well as lack of demand for monetization offerings are discussed.

Threat of weakening position of trust

For majority of the companies with valuable data, the evaluated risks seem often bigger than the potential benefits considering data monetization. Interviewed companies with valuable and unique data assets, for example in finance, telecommunications, health care and retail industry, were careful to take advantage of their valuable data. As one health care company pointed out, *"The position of trust we have with our customers is so strong, that the potential deals of monetization are not even considered, since their business value could not be worth of the risks in decreased trust position."* The issue of trust was an acknowledged trade off and possible the biggest barrier considering data monetization in majority of companies operating with consumers: The potential weakening related to trust and integrity are superior when compared to the potential business advantage gained from

monetization. Thus, the plans for monetization stay narrow as well, as one CDO noted: *"Currently we do not monetize, we cannot monetize nor we want to monetize data by identifying customers and selling this data to third parties"*. One director mentioned: *"Data is a matter of honor for us, so we do not want to test the boundaries of our customers."* The risks related to monetization can easily prevent creation of monetization solutions.

Small evaluated monetary benefits

Unknown potentiality of monetization slowed down investments, especially when the aforementioned risks were acknowledged as well. The monetary returns and viability of data monetization were raised by multiple companies. It was pointed out, that by sharing data or data derived solutions, the company has to financially benefit enough from it. Due to the novelty of business, it is hard to evaluate the benefits: The potential scalability as well as the pricing of monetization offerings were unclear. This was further underlined with the unknown potentiality and the challenges in creating calculations and business cases considering data monetization. Due to the novelty of business, the management's support is often needed to justify investment decisions in fairly risky monetization. Some companies underlined the niche market of data monetization and stated it can never provide sufficient revenue for established companies. Additionally, the data might have value for other companies, but it might not hold enough value for commercialization: The solutions need to be scalable enough to provide sufficient revenue when considered the costs of data monetization. The incoming revenue need to be sufficient to cover the lost advantage of shared data. In some companies, the clear incentive to create data monetization solutions was not evident. One director described the situation: *"We still lack the clear drivers and motivation: Why should we sell or monetize our data?"* The small evaluated benefits of monetization were a clear barrier, especially when combined with identified risks.

Lack of customers' demand for monetization offerings

The low demand for monetization offerings does not ease the situation. The lack of demand and customers' lack of openness for new solutions were mentioned in many occasions. Especially companies with data monetization solutions addressed the issue of creating demand: Customers need to believe that the data monetization solution is valuable for them and they need it as well. Proving the credibility and value of solutions can be challenging and take time, therefore the solutions need to truly provide value for customers and be faultless. A third of companies pointed out, that in addition to the requirement for data to create value to the customers, the customers

themselves have to be ready for the solutions. The maturity of customers was an issue present in all industries, except health care, where the utilization of data between different actors was scarce. As one media company noted: *"We grow and scale our current offering at the same rate as the markets and customers mature for data derived solutions"*. The high maturity of data premises in a company cannot be leveraged properly, if the potential customers are not willing to buy the solutions nor understand their value. Customers rarely ask for new data driven solutions nor do they offer them themselves. The unawareness of value of data and lack of customers' demand slows down the initiation of data monetization projects.

Following *Table 4.4* evaluates the criticality and frequency of aforementioned barriers for companies, based on the findings. Criticality refers to the gravity of issue: How significant the issue was as preventing data monetization. Frequency relates to the popularity of issue for interviewed companies. Both criticality and frequency are evaluated as low, medium or high.

Table 4.4: Barriers of monetization related to benefits and risks of monetization.

Issue	Description	Criticality of issue	Frequency amongst companies
Threat of weakening position of trust	The potential monetization can compromise the trust position company possesses. This poses a significant threat especially for companies operating with consumers, such as health care, finance, telecommunications and retail companies.	High	Medium
Small evaluated monetary benefits	The benefits of monetization are hard to evaluate, yet companies consider them small and lack clear incentives. Data monetization solutions need to be scaled to multiple companies to provide sufficient revenue, but the scalability of solutions is unknown.	High	Medium
Lack of customers' demand for monetization offerings	The customers' low demand for monetization offerings slow down the development of monetization offerings. As the demand is low, companies need to prove the value and credibility of monetization.	Low	Medium

4.4.2 Other data related projects are prioritized before data monetization

Other data related projects in companies affect the potential investments in data monetization. In this subsection, the prioritization of internal data projects as well as the focus on current customers are discussed.

Prioritization of internal data projects

It was noted by multiple companies, that the opportunity costs related to data initiatives delayed the experimenting with new offerings. As one CDO pointed out, *"some other investment can bring quicker results, than a seemingly risky data utilization project".* The lack of resources made it more difficult to create new data derived solutions, as new data derived service initiatives were not prioritized. Currently the potential benefit was higher in internal processes and data projects, as one telecommunications company noted: *"Currently the potential business cases in data utilization are better for internal processes and current products, than for data monetization."* The less risky internal projects provided a quicker, more justified and safer returns than unknown monetization projects.

Prioritization of customers through current products and services

The focus on current customers prevents creation of new offerings, especially for new segments. For majority of companies, the focus of data usage was on current products and customer relationships, not on new offerings, such as monetization solutions. As one CDO underlined the issue: *"The focus of data utilization is to create value to our customers, especially the current ones".* One director stated *"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".* The reasoning for providing data derived offerings for customers was not the potential economical value through monetization, but the additional value created for current customers. The aforementioned economical potentiality of monetization was underlined by multiple companies: Often the revenue and profit for the company comes from the core business and not from the monetization practices. Therefore, the focus of data utilization is often on current products and data monetization is one way to provide new value and differentiation to current customers. This was evident especially for companies operating with consumers: It was underlined that the monetization and cashing in with data is not any area of focus and is ethically suspicious as well. The data monetization solutions were seen possible, if the data was highly aggregated or the current customers received value from the monetization as well. Companies, who were more mature considering their data utilization, underlined the prioritization for their data utilization, and potential data monetization, with following questions:

1. How do our customers benefit from the utilization of data?
2. How do our internal processes, organization or decision-making benefit from it?
3. How does some other stakeholder benefit from our data?

Therefore, the gains of data usage were first considered related to current customers and internal usage, before creating any data monetization solutions for third parties. As one director brought up the point as well: *"We do not sell anything derived from our data. We have succeeded to create value to our customers from the collected data, which has been the first challenge."* The importance of customer centrality was emphasized by multiple companies, where the use of data was driven by the potential benefits for current customers. This affected data monetization as well, as the monetization offerings were less supported, as the focus was on current customers.

Table 4.5 evaluates the criticality and frequency of aforementioned barriers for companies, based on the findings.

Table 4.5: Barriers of monetization related to prioritization of other data projects.

Issue	**Description**	**Criticality of issue**	**Frequency amongst companies**
Prioritization of internal projects	Internal data related projects can be prioritized before data monetization. The projects might provide less risky and foreseeable returns than unknown monetization.	Medium	Medium
Prioritization of customer projects	The data utilization focused on the current customers. This prevents the creation of new data monetization solutions for new actors, as the monetization offerings need to create value to current customers.	High	Medium

4.4.3 Organization's culture nor capabilities are not suitable for monetization

The organization affects naturally the possibilities, and barriers, related to data monetization. In this subsection, the organization's culture regarding risk taking and data sharing and required management's support and competencies of organizations are discussed.

Experimentations and risks avoided

Company's culture can prevent data monetization by hindering new business creation, risk taking and data utilization. As companies without data monetization solutions noted, the organization's inability to notice business opportunities can hinder monetization: Companies with a long history and legacy faced the situation, where the company culture do not support the discovery of problems and finding solutions for them. A fourth of companies argued, that the lack of creativity slows down companies. Still, the stiff culture associated with minimal risk-taking was a bigger barrier for the creation of new business models, such as novel data monetization solutions. This was more common for manufacturing and other B2B companies, which had low maturity of data monetization. It is hard to come up with new services, when employees do not understand for whom the collected data could be valuable nor want to take risks to experiment new things. Therefore, the culture prevents the creation of new businesses.

Data is not shared outside the organization

In some cases, as few companies pointed out, the culture itself does not support the utilization of data. In multiple manufacturing and construction companies, the mentality of company and industry was seen as an issue for data utilization, sharing and eventually monetization: Companies do not easily share their data with others. One interviewee from a manufacturing company described the atmosphere of the industry as *"there is a dominant mentality, that we can manage on our own and do not need help from others, nor they need help from us"*. However, the culture and leadership need to support data utilization and its monetization. As one CIO noted: *"If there is no culture nor encouragement to share information even between different projects, it is hard to monetize or even utilize company-wide data properly."* To monetize data, companies need to first learn how to utilize it and share their data.

Lack of support from management

The focus and support needed for from management was addressed in all industries involved. One CIO even mentioned, that *"the support of management is the most important driver of data utilization"*. Thus to develop monetization solutions, the sufficient support and focus from the management of company is needed. Traditional IT was not seen as a sufficient basis for data monetization solutions, but experimenting and lean and agile approaches were underlined by multiple companies, especially in traditional industries. One media company described a suitable basis for monetization as *"A sufficient balance with enthusiasm, experimenting and support pro-*

cesses." This was underlined by the new venture accelerators, collaboration with start-ups as well as different pilot projects underway. Still, the need for data monetization to integrate with current business offerings was evident. Companies from different industries emphasized the need to locate data solutions closely to current business functions. The company itself need to be able to benefit from data solutions internally, and not just the external actors through monetization. Since data needs to be verified and described, and the optimization and pricing of final solution takes significant amounts of time, data monetization requires time, experimentations and failures. Organization's IT and management need to be suitable for data monetization to succeed.

Lack of skillful employees

The lack of competence and skills required for data monetization was brought up by about half of the companies. This was often realized by companies who were mature enough to initiate data utilization and monetization solutions. The tools and resources were previously seen as barriers for wide data refinement, but they were now seen more as enablers and drivers of data utilization. For most of the companies with insufficient talents, the problem lied in the amount of data scientists and analysts. *"We have skillful data professionals, but the bench is not wide enough"*, noted one director. The technological superiority, created by data scientists, was desired as well, but for most of the companies, the skillful, business understanding analytics were needed the most. Few companies underlined the pioneering know-how regarding technology, but acknowledged the shortage of commercialization and business professionals. Only one manufacturing company out of interviewed 19 companies had clear restrictions considering technology. In this particular industry, the current technology did not yet support the collection of valuable usage data. Therefore, the current data did not hold much value to other actors.

Table 4.6 evaluates the criticality and frequency of aforementioned barriers for companies, based on the findings.

Table 4.6: Barriers of monetization related to organization's culture.

Issue	**Description**	**Criticality of issue**	**Frequency amongst companies**
Experimentations and risks avoided	The company culture does not support risk taking, which slows down the creation of new services. Experimenting, discovering of problems and creating solutions for them are not performed.	Medium	Medium
Data is not shared outside the organization	The company culture does not support sharing of data. Therefore, data is not shared outside nor inside of the company, which prevents creation of monetization solutions.	Medium	Low
Lack of support from management	Data monetization requires support and focus from the management: As the returns for data investment are hard to calculate, enough time and resources, as well as space for experimentations and failures are needed for monetization to succeed.	High	Low
Lack of skillful employees	To produce viable data monetization solutions for customers, the premise for data monetization need to be in order. The competition on skillful employees regarding data science and analytics is intense, and company can lack enough talent to produce desired solutions.	Medium	Low

4.4.4 Data is not easily accessible nor in good shape

For most of the companies, the data posed as one of the most significant issues of data monetization. The issues considering data related to its quality, accessibility, ownership and regulations.

Insufficient data quality

About half of the companies saw data quality as one of the biggest issues affecting data monetization. Data has to be in the right and accurate form for successful utilization, and this was not the reality majority of companies faced. The bigger the company was, the more challenging was the quality of data for the company. The challenges regarding quality were acknowledged, as one media company remarked: *"If data is created as a side product in manufacturing and traditional industries, it can be extremely challenging to monetize it. Our main product is data, and still we have had hard time to ensure the sufficient quality of data to monetizing it."* In some industries, data can be originally in a very non-structured form, as was the case for health care companies. The data can be stored as open form text, which is not connected to the structure of data and its variables, which makes the monetization harder. In addition to data being in a wrong form, the richness of data can be a issue. If the data itself does not contain multiple variables or is not rich in details, the seemingly good quality of data does not help the monetization. Richness of data cannot be evaluated without context, as it depends on the use case of data as well. This was addressed by one CDO, who noted that *"the accuracy of data is not valuable intrinsically, as it depends on the use case. Too accurate data with multiple data points might cause additional complexity and problems considering data storing, if the value is derived on more abstract level."* B2C companies as well as media companies saw quality of data as a vital competitive advantage, since they had done significant investments considering the quality of their data. It was acknowledged by almost all companies, that it takes commitment, time and competence to ensure high quality of data and to prevent it to become corrupt. As some companies pointed out, the quality of data and the possibilities to enrich it require continuous development, especially with development of supporting processes and infrastructure around them.

Weak access to data

In addition to data quality, accessibility to data was acknowledged as a significant issue. If the data is not easily available in the company, it is not possible to share it properly. Therefore, it is not enough for data to be in a right form, if it cannot be accessed easily with used tools. This was evident for companies in all industries: The internal silos hindered the utilization of

data, or had been one of the biggest barriers for more mature companies. The silos were often created by the lack of common practices, by the vast amount of fairly autonomous business functions, mergers and acquisitions of other companies and insufficient integration of different business functions. In many cases the issue was due to data generated and located in different business functions, which did not interact and communicate with each others sufficiently. The impact of legacy systems, acquisitions of new companies and dozens of used IT systems all contributed to increasing barriers between different data silos. In addition, the consolidation of data was noted to be troublesome, since the data, its richness by amount of variables, structure and used time frame can differ in all used systems. Therefore, the integration of different data silos remained to be a clear barrier. However, some companies had successful development around the issue, and especially media companies had overcome the issue of challenging accessibility of data. The harmonization and consolidation of data into a sufficient master data management system was found as a significant investment ongoing in multiple companies. Value of data, as well as the output of monetization, depend how relevant and up to date the data is. The value of data and monetization often do not increase, as the data grows older. As one media company mentioned, *"Companies tend to preserve data in their storages. Still, the value rarely increases when data is stored."*. Companies which had already monetized their data, addressed the issue of external data integrations as well. Data has to integrate with customers' systems as well, but the retrieval of data from internal systems in a right form is more essential.

Ownership of data

In addition to quality and accessibility, data ownership raised issues in almost half of the companies. The data ownership requires negotiations and careful approach as every company aims to own the data themselves. In most companies the data ownership varied, so that the company and its customers owned the data depending on the use case. Nevertheless, the ambition was for all interviewed companies to keep the ownership of data themselves. The issues with ownership were most evident for manufacturing companies and other B2B companies, as well as health care companies. As one CIO pointed out, there are *"significant contradictions considering interests and ownership of data when it is potentially monetized"*. This relates to issues, such as Who is the central actor, who originally owns the data and who gets the biggest share of the generated revenue. Few companies tackled the issue of ownership by not trying to own the data, but negotiating how they can utilize the customers' owned data. Still, these companies acknowledged that when the data is owned by the customer, there are less opportunities to utilize it.

Table 4.7 evaluates the criticality and frequency of aforementioned barriers for companies, based on the findings.

Table 4.7: Barriers of monetization related to organization's data.

Issue	Description	Criticality of issue	Frequency amongst companies
Insufficient data quality	Data need to be in a right form to sell it further. It need to be rich in details, accurate, continuous and complete as well.	High	Medium
Weak access to data	Data need to be easily accessible for proper utilization. The consolidation of different data sources can easily cause issues, as the data is collected from different systems and business functions.	High	Medium
Ownership of data	Data ownership can act as a barrier for monetization, as it is hard to monetize data that companies do not own. The data ownership creates additional friction in monetization, as companies try to get the ownership of data in ecosystems.	Medium	Low

4.4.5 Privacy and legal issues cause careful approach

Different industry-specific regulations, consumer focused regulations, privacy and security act as potential issues companies need to consider when monetizing data.

Strict industry specific regulations

Regulations and legislation were seen as evident factors restricting the potential monetization by majority of companies, in total 14 out of 19 interviewed. For most companies, the Finnish legislation prevented the direct selling or utilization of customer data for monetization purposes. For media

companies, the privacy protection complicated the data monetization, as the selling of data is restricted. For some companies, for example for insurance companies, the legislation restricts performing any other business than insurance. Therefore, selling of the collected data as a new business was not possible. Few companies stated, that the legislation in Finland was too slow to change, as it acknowledges privacy poorly and restricts the possibilities for modern data utilization.

GDPR causes carefulness

The changing legislation environment increased the carefulness of companies in multiple industries. Especially for companies operating with consumers, the presence of General Data Protection Regulation (GDPR) by the EU Parliament was ubiquitous. Majority of companies had prepared themselves for the change, but saw it still as a challenge and a significant reason for careful approach on data monetization. One media company described the situation: *"GDPR slows down investments related to data. The lack of precedents especially increase caution since no one wants to be the first to test the boundaries."* However, as few companies pointed out, GDPR could open new monetization opportunities, since customers can start monetizing their own data easier. The investments and work required due to GDRP were seen significant, regardless of the size of company. Still majority of companies affected by GDPR, underlined the increased transparency and shared common rules as a good thing for data markets.

Secure solutions are required

The issue of information security was mentioned by few companies, but it was not seen as a barrier preventing potential monetization. Security of data solutions was addressed as an important issue, but seen more as a hygiene factor to consider and to take care of, not as a continuous barrier preventing data monetization. Nevertheless, as companies engage more in data monetization, the role of security might increase, as companies face different practical issues, especially when monetizing data considering individuals. Currently the barriers companies face relate to creation of monetization solutions. Yet, security is more related to operational issues of monetization, which is still an abstract situation for most of the companies.

Table 4.8 evaluates the criticality and frequency of aforementioned barriers for companies, based on the findings.

Table 4.8: Barriers of monetization related to legal and security issues.

Issue	Description	Criticality of issue	Frequency amongst companies
Strict legislation in data monetization	Legislation sets different limitations for companies, as it can restrict the utilization of customer data for monetization purposes. This relates for example to finance and health care industries, where legislation in Finland and in EU prevents the broad utilization of data and often the direct monetization of data.	High	Low
GDPR causes carefulness	GDPR causes companies to approach consumer related data carefully, as the potential penalties are significant. In addition, investments due to GDPR slow down investments to other data related projects, such as monetization.	Low	Medium
Secure solutions are required	The security of monetization solutions is not a barrier for companies. Still, companies need to ensure a sufficient security, especially when consumer data is considered.	Low	Low

4.4.6 Variety of barriers slow down the creation of monetization solutions

The barriers companies face are often related to industry and the current state of data utilization in the company. The biggest barriers for companies are the following: As the evaluated monetary gains of monetization are yet unknown or small for companies, the benefits do not outweigh the acknowledged risks, such as weakening in position of trust. Therefore, often other data related projects are prioritized before data monetization, as they contain less risks and can provide more foreseeable returns. For some companies, the current culture or capabilities slow down the potential monetization. Especially the lack of management's support can be a significant barrier. In addition, the quality of data and easy access to it are one of the biggest barriers, regardless of the industry. Legislation causes companies to approach monetization carefully, and acts as a clear barrier in some industries. The variety and joint effect of different barriers create a challenging premise for data monetization in Finnish companies.

Chapter 5

Discussion

This chapter combines the results from interviews with the academic literature presented in *Chapter 2*. The discussion involves a deeper analysis of results, while discussing the relationship of the results to the reviewed theory. This thesis focused on data monetization, how companies currently monetize their data and which issues affect this new monetization of data. Following sections answer the three research questions used in this thesis, as well as provide recommendations, practical and theoretical implications and thoughts for future research and limitations. Finally a conclusion is provided to summarize this thesis.

5.1 Answers to the research questions

RQ1: What does data monetization entail for firms today?

The first research question concerns the concept of data monetization. Data monetization is using data possessed by the company to create new revenue. The revenue is created by providing a distinct service or as a part of current offerings. The monetization can be realized by providing data or data derived offerings to current customers or new ones. The question RQ1 was discussed more thoroughly in *Chapter 2*: Data monetization is a novel phenomenon in academic literature and still lacks a clear definition: Fred (2017, p. 24) described the term as *"the revenue generation with and out of data and data-derived and information-based products and services"* in her literature review regarding data monetization. This definition does not acknowledge the differences between customers of data monetization: The monetization is significantly different process, if the customer is a current customer, who receives the value as an additional feature derived from data,

or if the customer is a new customer. This comes to the nature of data monetization as well: It can be seen as an external service reusing gathered data, or it can be implemented as an addition on top of the current relationship with customers, partners and other actors. Therefore, the monetization can be implemented as a part of current product or service, or it can act as an own service for new customers. This makes the monetization broader, when the focus can be on current customer relationship, where data sharing and implementing data derived features into current relationships are considered as data monetization. Yet, there is always at least two actors involved: Actor monetizing owned data and the actor using the data. The academic literature lacks this kind of discussion considering data monetization, when the phenomenon is not strongly defined. This thesis summarized the definitions and theory considering data monetization from multiple sources, and provides a more comprehensive definition for the term.

Data monetization: Practice of using possessed data to create monetary value. At least two actors are involved: Actor monetizing the possessed data and the actor using the data. The value of data is provided by a distinct data-derived service or as a part of current offerings. The practice is implemented by providing data or data derived offerings to another actor; current customers or new ones.

RQ2: How do firms monetize their data?

The second research question concerns the practice of data monetization. Data monetization is industry, company and use case driven, which affect the monetization. Companies can create different offerings, such as providing raw or aggregated data, providing insights or analysis as well as creating a scalable data derived service. All of these offerings can be combined and provided to current customers as an additional service, as well as a distinct service open for everyone. Currently, the focus is more on the current customers than other actors.

RQ3: What affects data monetization done by firms?

The third research question concerns the issues affecting the data monetization. As in second research question, the issues depend a lot of context. The changing business environment with big data, cloud computing, analytical tools, changing legislation as well as market pressure enable and accelerate the further utilization and monetization of data. Still, the quality of data and accessibility to it, threat of decreasing position of trust, unknown monetary benefits, organization capabilities and culture as well as legislation slow

down the creation of monetization offerings.

The research question RQ1 acted as a prerequisite and basis for questions RQ2 and RQ3, as the focus is on the latter questions. The following sections provides a deeper level discussion considering the findings of this thesis and the research questions RQ2 and RQ3.

5.2 Current state of data monetization

The data monetization is a topical phenomenon, as there are a lot of enablers and drivers that have not been present earlier: Big data masses with supporting trends and technologies, such as cloud computing and better analysis tools and data capabilities, enable companies to utilize new opportunities regarding their data. The development around aforementioned topics accelerate the broader utilization of data, in which data monetization is one way to utilize data more broadly. The changing legislation further supports monetization as PSD2 encourages financial companies to monetize data, while GDPR requires investments to data infrastructure. However, the market pressure and the threat of disruption accelerate the utilization of data in different companies: The better capabilities enable monetization, but the motivation and incentive often come from markets. As Amit and Zott (2012) and Vives (2008) note, competitive pressure fosters innovation and accelerates the higher prioritization of new solutions. Monetization with identified and unknown risks is rarely experimented, if there is no clear driver or pressure behind the reasoning, especially when there are less risky internal data related projects available. In this way, the companies can be divided to three categories depending on the market pressure for broader data utilization and new innovations:

- **Companies before data disruption**: In industries, such as manufacturing, construction, electricity and health care, companies lack a clear market pressure and incentives to seek for broader, risky ways to utilize their data. Then companies can still focus on internal projects, and monetization as a fairly risky way to utilize data is avoided. Therefore, there are no evident incentives to start projects related to novel data monetization.

- **Companies experiencing data disruption**: Some companies are forced to rethink their data utilization. Finance sector is facing pressure to evolve into more digital, data utilizing industry due to PSD2. The threat of novel fintech players as well as competitors utilizing com-

pany's data can encourage previously careful industry to pioneer in external data utilization. Pressure for a broader utilization of data can also be accelerated by other factors than legislation, such as decreasing revenues. This is the case in forest and logistics industries, where companies are seeking out new revenue sources to survive in changing markets.

- **Companies after data disruption**: Media industry has experienced already a disruption with the transformation from traditional media to digital and data-driven one. During the disruption, companies have been forced to utilize their data to survive and data monetization has been one potential solution to capitalize on the data assets. In addition, the increasing pressure from losing market share of digital advertisement to global actors such as Facebook and Google accelerate the need to monetize current data assets.

Despite the aforementioned enablers and drivers of monetization, companies do not monetize their data broadly. There are some comprehensive monetization solutions, yet the majority of companies do not engage in monetization. A significant amount of companies are laggards in their data utilization: It is hard to monetize data, if the company has not yet succeeded to utilize it internally. As Najjar and Kettinger (2013) suggest, companies should first establish sufficient capabilities to address internal business needs before sharing or monetizing data with parters: Companies can achieve easier, foreseeable and quicker results by investing on internal processes and current products, not in creation of new data derived solutions. If there is no evident pressure or threat of disruption, the smaller amount of stakeholders, more visible benefits and easier controllability of internal projects all contribute to companies first focusing on internal data projects. Majority of the Fortune 1000 companies focused on decreasing expenses with their big data initiatives, not on creating new revenue (Bean, 2017).

The data monetization, while being an unknown business, can provide companies new ways to improve current relationships and differentiate company's offerings. This is a popular way to monetize data, as an additional data derived service can personalize and improve current products and offerings. Thus, monetization can be used as a way to increase customer-centricity while creating new value to all parties involved. New services often originated from collaborations with single partners, since companies initially experiment with customers and partners, as Thomas and Leiponen (2016) suggest. It is natural for companies to start sharing more data with the current partners, due to improvement of relationships, more foreseeable value of data to

other parties as well as more controllable flow of data. Despite the partnership driven monetization, the most pioneering companies regarding data monetization focus on a scalable approach. The aim of monetization is often to provide a solution for multiple customers to ensure the economic viability, yet the initial monetization can be achieved easier through partnerships. Therefore, the data monetization solutions can be first tested and validated with one actor, before a possible broader scaling. However, the monetization can already create a competitive advantage for the company through differentiated offering.

As Spijker (2014) and Thomas and Leiponen (2016) suggest, the data monetization is not limited to specific business models or offerings, but is often integrated into current relationships. The monetization solutions can be divided to specific offerings, for example to selling of data, providing insights or analyses and creating scalable services or products as presented in *Chapter 2*. Due to integration of monetization solutions to current business, the offerings can be part of current relationship with customer or can be provided to other actors. The following framework, *Figure 5.1*, summarizes the ways to monetize data, based on this research.

Amount of customers →

Level of refinement ↓

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	Sell aggregated data considering end-users to current suppliers	Sell aggregated data on market activity to investors and authorities
Providing insights	Provide insights to current customers considering their business environment	Provide trend and demand insights to suppliers	Provide analysis of consumer demand to investors
Creating a scalable service	Provide a service, where customers receive information of business environment	Provide a service, where suppliers can analyze end-user consumption information	Provide a service, where investors can access the real-time information considering market trends

Figure 5.1: Options for data monetization.

The refinement level, as well as the potential amount of revenue, increases as companies move from selling data towards a scalable service. The potential revenue increases as well when the potential amount of customers increase: When a service can be provided to partners as well as current customers, the revenue is higher when compared to service provided only to single actors.

Monetizing company's data is not a simple thing to do, as it is affected by multiple different issues such as monetary attractiveness, organization's culture, data management as well as legislation:

- Currently the markets of data monetization are niche markets, where the scalability and wider adaptation of solutions are fairly unknown. The position of trust established with customers is a valuable asset, which can be compromised by monetizing customer data. The uncertain economical potentiality, risks involved as well as the premise for monetization weakens the attractiveness of monetization, as it is easier to focus on internal possibilities to utilize data than monetization.

- The culture is one of the biggest barriers for further utilization of data, as the culture can prevent experimenting, risk taking and trusting to other companies. Data monetization is uncharted territory, which raises uncertainty with unknown demand as well as with the risks involved in sharing the valuable data outside the company.

- In addition to risk evasive culture, the data infrastructure is not mature enough for a great deal of companies to support the monetization of data: Different silos, legacy systems and acquisitions prevented the easy utilization of data. The data quality and easy access to it are often the first barriers to overcome, as Spijker (2014) suggested.

- For companies operating with consumers, the legislative restrictions can be strict considering the utilization and sharing of customer data. The changing legislation through GDPR delay monetization solutions further for most companies: Due to potentially high penalties involved, no company aims to try the boundaries and act as a precedent for GDPR violations.

The role of monetization is still very small for majority of companies, yet it can provide possibilities to differentiate and strengthen current relationships. The role of monetization for company depends greatly on the goal of data utilization: Does the company aim to create new value for current customers, to broaden the portfolio of offerings, to increase revenue or to improve valuation by providing new digital services? The goal of monetization affects the role of monetization, as it justifies the required investments

and focus for the service. Currently the creation of new customer value is the most popular goal, as the monetary returns of monetization of are still indeterminate for most companies. The ambition for monetization is needed for monetization solutions to succeed, as the barriers and investments are not easy to ignore, while the markets can be considered niche.

5.3 How to approach data monetization based on findings?

As companies have a varying level of data utilization, there are no general rules how to monetize one's data. However, some general guidelines can be drawn from the research.

Data-based solutions should be not be built on an unstable premise, which results in need for proper data infrastructure as well as a sufficient culture to support monetization. Companies need to invest in their data infrastructure, which means high-quality data, which is easily accessible and secure, while attracting enough capable employees to utilize and refine the data. A better data infrastructure, for example with APIs, can provide benefits already inside the organization by improved agile utilization of data, increased sales as well as cost savings (Benzell et al., 2016). Before monetizing data, companies need to understand their current position and goals considering data monetization (Najjar and Kettinger, 2013), as well as the current capabilities, data and different monetization possibilities (Laney et al., 2015; Fred, 2017). As companies develop a sufficient premise for data monetization, they can approach data monetization. The potential customers for data monetization can be easiest found close to the monetizing company: It is probable that current customers and companies in value chain or value network appreciate the value derived from data. A strategical partnership ensures a more customized relationship, where the potential value of data can be explored more flexibly. Since companies need to validate the potential value of data first before actual monetization, there is a need for collaborations with other actors (Spijker, 2014). By iterating with customers, companies can easier validate and achieve the value creating data monetization solution.

These changes require significant support and strategical focus from company's management, as these changes require investments, time and change management. Current IT processes are not often suitable for data monetization solutions, since the unknown business environment is recommended to approach by experimenting. as companies can already struggle in creating viable business cases for current data initiatives, trying to reason the po-

tential benefits of monetization compared to risks can be challenging with traditional tools. Therefore, data monetization need to be approached with a more risk-tolerant strategy, implemented with agile experimentations such as parallel startups, corporate ventures or with startup collaboration. By establishing a separate subsidiary, companies can tackle potential issues regarding mother company's position of trust, core business, lack of agility and culture, as well as legislation in some cases. By establishing or acquiring another company, the initial brand and trust are less affected by the data monetization solutions. This is a popular way to experiment new data-based businesses: Insurance company Allstate span out a startup Ariety to monetize the data collected by customers (Kokalitcheva, 2016), while another insurance company Baloise established a startup to provide smart data-driven insurances for cars (Baloise Group, 2017) and Michelin established Michelin Solutions to provide new data-driven innovative services (Michelin, 2017). The new startups can easier build its business and provide services to the competitors as well, while operating under different industry and regulations. These experimentations ensure the required room for experimenting and iterations, as there is no direct pressure to provide continuous results nor the unit is limited to company's current culture and processes. Thus, autonomy, management's support as well as emotional commitment need to be ensured, as Mackenzie et al. (2014) and Spijker (2014) propose.

Companies can approach data monetization with following steps:

1. **Understanding the possessed data and capabilities**: Company can utilize its position in value network's intersections by monetizing data by collecting it from one side and providing it to another. For example, company's suppliers can value information from customers, or customers value data considering other customers. Still companies need to identify their current maturity and goals considering data capabilities before evaluating monetization possibilities.

2. **Discussions with suitable partners**: Companies can discuss with different companies to understand better the potentiality of business opportunity of their data. The suitable companies can often be found in current value chain or value network. The discussions should include actors outside the value chain as well, since often the common end-customer is enough for the data to be valuable. For example, authorities and investors as well as other data-driven actors might value the data.

3. **The quality of data and easy access to it are ensured**: Since value of data monetization depends on the data itself, the easy access to data as well as the sufficient quality of data should be ensured. The proper data infrastructure is required for ongoing data monetization.

4. **A subsidiary or a separate venture is established to develop monetization**: Using a subsidiary or a distinct team can result in more flexible experimentation. Agile development is easier, initial company culture is not slowing down and position of trust and initial brand are less affected by the experimentations through new brand.

5. **The value of data is validated with a chosen partner**: The promised value of data and the data derived solution should be tested and proved to customer before creating a comprehensive monetization solution. This can be implemented with pilot project or similar experiment, where the customers, as well as the company, get initial results of the value of data. This helps pricing and optimizing the data as well as communicating the need and value for customers.

6. **The increased amount of paying customers**: Start increasing the amount of monetization customers by selecting chosen companies for a limited and controllable data monetization, or create a scalable service open for all companies. The data sharing can be first started with a limited amount of current partners, who appreciate additional information on their value network or the common customers. After experimenting and optimizing service with limited partners, it can be easier scaled for other actors as well.

7. **Processes and roles support the continuous monetization**: Cost-effective and high-quality monetization can be ensured by establishing roles and processes to support ongoing monetization. The continuous generation and sharing of high-quality data is vital for long-term monetization.

These steps can be used to approach the practice of monetization. However, as noted in the earlier section, the monetization is often integrated to current offerings and depend on the chosen strategy. Therefore, companies need to choose a suitable offering for the monetization. The following figure, *Figure 5.2*, map the potential paths for monetization solutions.

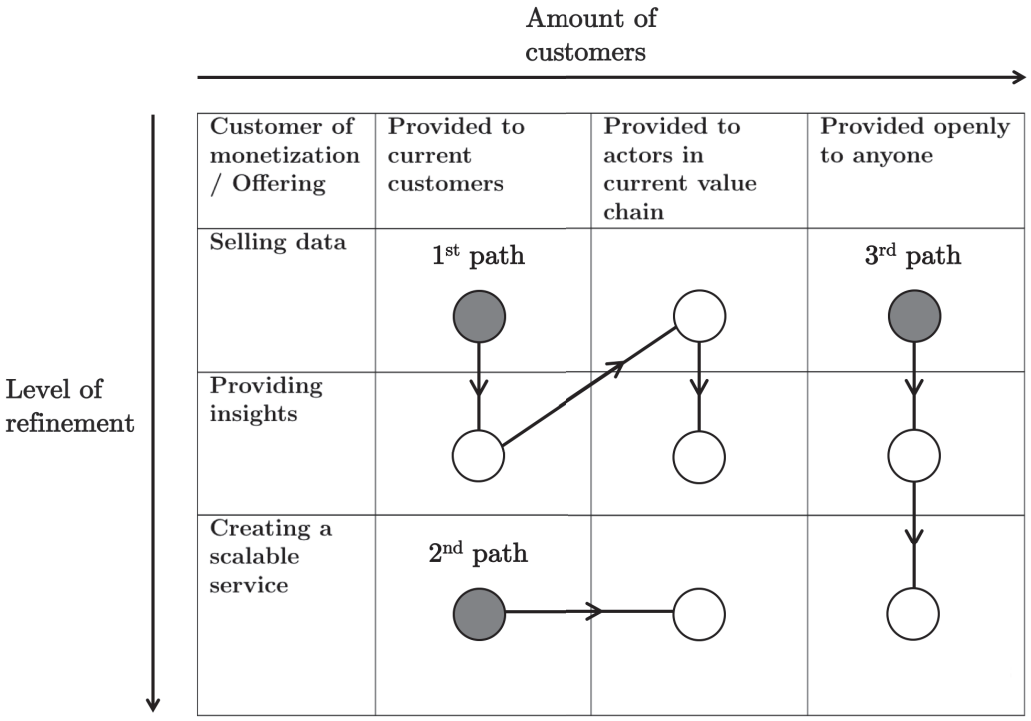


Figure 5.2: Different paths to approach data monetization.

The easiest way is to sell fairly unrefined data to current customers and partners. However, for multiple companies the step from not sharing data, to selling data it to multiple companies is a huge one. Therefore, companies can start monetizing their data first by providing it to their customers and partners. The first path of data monetization is about validating the value of data for other actors. There it can be provided to current customers and partners without deeper refinement. The second path still provides value to current customers and current value chain, as it is easier to control offerings provided to limited amount of partners. Still the level of refinement is higher, as the data is no longer provided as datasets or insights. When monetization solutions have been provided to current customers and actors in the company's value chain, they can be provided openly to any actor if needed. This way, new customers from different industries can be attracted, yet this third path should be the last one to be provided due to higher risks and investments involved in it.

5.4 Practical implications

The practical objective of this thesis was to provide information and coherent understanding of data monetization in Finnish industry. The goal was to understand how big Finnish companies currently monetize their data and which issues affect and prevent the potential monetization. In addition, the aim was to provide recommendations how companies could approach the novel monetization. The practical contribution of this thesis is a better understanding of the data monetization: The provided theory with examples, new definition of monetization and a suitable framework and recommendations based on theory and findings, all contribute to a more comprehensive understanding of data monetization.

Data monetization is affected by multiple different things, such as company's capabilities, other projects, culture, the surrounding business environment and markets as well as legislation. Thus, it is not a simple practice, as there are a lot of different issues affecting the attractiveness of monetization as well as the possibilities in it. This explains to some extent, why it is not common practice, since there are a lot of issues affecting it. In addition, it has not been possible before, due to technological capabilities.

Data monetization is still a nascent business for most of the companies. However, there are different monetization experimentations and comprehensive solutions offered in different industries. The provided framework presents the identified options for monetization: Monetization solutions can be offered to current customers, new actors in value chain or openly to any actor by selling data, providing insights or creating a scalable service. For most companies, data monetization is still a slightly faraway goal, as companies are not yet suitable for data monetization nor is the monetary attractiveness of monetization sufficient for companies. Companies with developed data infrastructure and capabilities, can create new business and improve current ones with data monetization. Majority of the monetizing companies focus on providing monetization solutions to current customers, while by creating on top of current products or services, monetization can provide personalization and differentiate offerings.

For companies considering data monetization, the first step is to understand the current situation with the data usage. It is not recommended to invest in risky businesses on unstable premise, so the current level of data infrastructure and maturity need to be understood. By understanding own capabilities, companies can approach the monetization by the most suitable way for their organization. Earlier section provided some practical recommendations how to approach the monetization. Data monetization can be

approached by creating monetization offerings on top of current offerings and relationships, as it is easiest to test and implement. However, companies can take a strategic approach to it as well, by creating the offerings without collaborating with current partners and customers. Still this requires a strong support from management as more investments and better capabilities are required. As there is no clear demand for monetization, and the markets appear as niche and unknown markets for companies, this is one possible way to approach the monetization. While companies improve their data infrastructure and achieve better data capabilities, they can utilize the possessed data better. Then the data monetization appears as a more attractive option, as the investments required for monetization are smaller, when companies can utilize their current capabilities and data infrastructure. In addition, while the popularity of monetization increases, the markets and demand of monetization are more foreseeable as well, which will ease the prioritization.

5.5 Theoretical implications and future research

The theoretical objective of this thesis was to provide information as well as on the novel topic of data monetization. The academic research is lagging on the field of data monetization, as the literature has focused mainly on internal usage of emerged big data, as noted by Fred (2017) and Thomas and Leiponen (2016). Currently there are no clear definitions of data monetization nor data commercialization, thus this research provides a comprehensive definition of data monetization.

Data monetization: Practice of using possessed data to create monetary value. At least two actors are involved: Actor monetizing the possessed data and the actor using the data. The value of data is provided by a distinct data-derived service or as a part of current offerings. The practice is implemented by providing data or data derived offerings to another actor; current customers or new ones.

The provided definition acts as a clearer definition as earlier ones, and covers the acknowledged issues affecting the monetization. It addresses the role of monetization as part of current offerings or as a distinct service, which had been ignored in earlier definitions. In addition, the customer of monetization affect the definition as well, as the monetization, its goals and practices are significantly different if the customer is a current customer or if the monetization is provided openly to anyone. By noting the presence of two actors, the

monetization is more defined in the borderline cases of internal and external monetization.

The theoretical contribution of this thesis is the further definition of data monetization: What is data monetization, what does it contain, and what is it like? This research contributes to the academic field by providing new empirical results regarding data monetization solutions, issues affecting them and the whole field of monetization in Finnish industries. Additionally, concrete recommendations and things to consider when approaching monetization are provided, which researchers and companies can take into consideration in the future studies and practices. This research provides new theoretical summary with global examples as well as a framework based on the theory and empirical findings to understand the monetization better.

For further research, there is a clear need for a more descriptive research on data monetization. This is further accelerated by increasing interest and experimentations by companies, as well as the rise of drivers for monetization. The research has focused mainly on big data and Internet of Things, as the external utilization of created data masses has been disregarded. Currently the academic literature does not reflect the reality, where companies struggle with their data sharing and commercializing initiatives. The proposed definition in this definition should be tested and validated further in the future research. As the approach of this research to data monetization was exploratory, a more interpretative approach with specific industries should be taken to explore the underlying factors. For example, the internal and external monetization could be studied further, as many actors have focused on internal monetization and regarded it as monetization. However, the borderline cases between internal and external monetization, such as improving current product features with data or adding new data derived features, are unclear in earlier definitions of monetization. Thus, the future research should first focus on the proposed definition of data monetization, followed by industry and issue specific researches. For example, in finance sector, the PSD2 will remove multiple barriers of data sharing and monetization, which will change the business environment radically.

Based on the provided framework, potential paths and the empirical analysis, the data monetization practices and business models should be studied further. Today companies approach monetization with varying practices, yet the current literature notes this variety of business models and offerings weakly. By researching more the successful monetization practices, the path to monetization and strategical monetization could be understood better: Do companies create successful monetization practices in collaboration with customers, on top of current offerings, or are they created as a separate strategic decisions, as the demand of customers is yet scarce? In the

future, academics can focus on monetization and build on the basis set on this definition, framework and the provided findings.

As there are new emerging business models, the field of data monetization offers great possibilities for future research considering data-based business models and how they relate to organization's strategy. As data monetization can be integrated to current offerings, or it can act as an own service, or even platform, the field provides a vast amount of potential research topics. The future research could focus on the possibilities of data monetization and platforms, as the same data can be utilized for multiple actors, while new actors provide new data to the platform. The results of this study suggest, that data monetization practices and business models should be covered more thoroughly in the academic research. Future research regarding the possibilities of monetization could focus on the possibilities to utilize the improved data infrastructure: While companies improve their data capabilities, the probability to engage in monetization increases, as it is easier, cheaper and less risks are involved. As companies improve and refine their data and data management, the investments related to monetization are smaller and attractiveness of monetization might increase.

5.6 Limitations

The strategical and exploratory approach of study aimed to create a comprehensive understanding on data monetization. Multiple case study approach, presented by Yin (2009), was used to collect and analyze the qualitative data. Therefore, four criteria presented by Guba and Lincoln (1989) are used to evaluate the validity of qualitative research: Credibility, transferability, dependability, and confirmability.

Credibility

Credibility discusses how truthful and believable the results are (Guba and Lincoln, 1989): Are the findings valid and about what they appear to be about? The used semi-structure interview reflect the opinions of interviewees: The data considering subjective issues is prone to credibility issues such as subject error and biases, since there was only one person interviewed in every company. To draw a comprehensive understanding of data monetization in researched companies and industries, interviews covering multiple roles per company should be conducted. Thus, especially when exploring the current monetization and faced issues, the interview data inevitably contained subjective assessment. The interviewer aimed to keep the discussion as close to company's reality as possible, yet interviewees described their

own business, so they can be prone to subjective observations. In addition, interviewees can promote their companies, for example by overemphasizing the importance of trust over additional revenue, which distorts the collected data.

Transferability

Transferability refers to the possibility to generalize the results to other settings or research subjects (Guba and Lincoln, 1989). The weak transferability of results is common for qualitative case studies (Miles and Huberman, 1994). It is evident for this thesis as well, that the findings are not generalizable for all other companies, which might operate in different industries and countries, where different regulations, culture and practices are faced. As often in qualitative research with case studies, the aim of this research was not to produce a generalizable theory to all populations (Yin, 2009), but findings that can be used to explain data monetization in a specific context. Due to high amount of different industries amongst researched companies, the transferability of this research is weaker than of a study focusing on a single industry. This was an acknowledged solution from the author: By focusing on a single industry, the exploratory nature of the research would have suffered. With a considerably high sample size of 19 different companies, the point of data saturation could be achieved within the studied phenomenon. As this thesis focuses on the data monetization in Finnish companies, it is acknowledged that most of the presented issues, for example GDPR or cultural rigidness affecting monetization, are broad phenomena and should be studied in-depth to provide more strategic implications.

Dependability

Dependability describes the consistency of the research (Guba and Lincoln, 1989): Did the data collection and analysis techniques yield consistent findings? The dependability was ensured by systematically analyzing observations from interviews and validating results against the initial data. The use of written notes as well as audio recordings increased the consistency of the study, as the results reflect the interviews. Due to presence of different sized companies from different industries, the sample size was required to be high enough to draw any descriptive findings. Still, the limited resources available to this thesis affected the sampling: The sample consisted of only single data points from some industries, such as construction and telecommunications, as well as companies with significant differences in their sizes, as some companies had yearly revenue over €3 billion and some less than €100 million. It was noted, that a senior data analyst might have more technical view on the phenomenon of monetization, whereas CDO might see the issues

from a more strategical point of view. The barriers of monetization were observed from different roles, which might have affected the prioritization and gravity of different faced issues, as people can see the closer and latest issues more significant than they really are (Saunders et al., 2009). However, due to relatively high sample size for a qualitative research, a single interviewee did not have a deciding effect on the results. The presence of multiple industries ensured that monetization could be explored more broadly.

Confirmability

Confirmability is about the neutrality, lack of biases, values and prejudice in the study (Guba and Lincoln, 1989). For interviews, a pre-reading was sent to interviewees before the interviews to define data monetization and ensure a focused discussion during the limited time. The pre-reading was not used in interviews per se, but might have guided the discussions and limited the discussion to revolve around previously defined concept of monetization. The pre-reading can be found as *Appendix B*. The semi-structured interview allowed the interviews to focus on possible monetization practices, but caused variance between different industries: If company had not monetized data, the discussion focused more on barriers, whereas with monetizing companies the focus was on current monetization. As any qualitative research, this study can consist of observer errors and biases (Saunders et al., 2009), still the interview structure, pre-reading for interviews, methodology and data analysis are presented transparently in this research, to ensure the confirmability of this research.

The scarce amount of literature considering data monetization limited this thesis to some extent: The opportunities and the depth of this thesis were limited due to lack of academic literature. Therefore, there are literature used from business domains as well as conference papers and similar articles, which affect the reliability and generalizability of findings. As a qualitative study, this does not try to draw generalizable results, but more to develop a deeper understanding of the researched phenomenon. There are other factors and limitations affecting the research as well the aforementioned ones, but the most critical ones were discussed in this section.

5.7 Conclusion

This research approached the concept of data monetization exploratorily. The object of this thesis was to increase knowledge about data monetization, and how it is practiced currently in Finnish companies.

Data monetization is a practice of using possessed data to create monetary value. At least two actors are involved: Actor monetizing the possessed data and the actor using the data. The value of data is provided by a distinct data-derived service or as a part of current offerings. The practice is implemented by providing data or data derived offerings to another actor; current customers or new ones. Data monetization is still a novel phenomenon in academic literature as well as for majority of companies, since companies rarely provide data derived solutions to other companies.

Data monetization is industry, company and use case driven, which affect the monetization. Companies can create different offerings, such as providing raw or aggregated data, providing insights or analysis as well as creating a scalable data derived service. All of the provided offerings can be combined and provided to current customers as an additional service, as well as a distinct service open for everyone. Currently, the companies focus more to provided value to the current customers than other actors. Yet, the pressure from competitors and disruptive market trends encourage companies in some industries to utilize their current data assets even broader and more aggressively. The changing business environment with big data, cloud computing, analytical tools, changing legislation as well as market pressure enable and accelerate the further utilization and monetization of data. Still the quality of data and accessibility to it, threat of decreasing position of trust, unknown monetary benefits, organization capabilities and culture as well as legislation slow down the creation of monetization offerings.

The data monetization can be approached with partnerships and increased data sharing with current customers and actors. It is important to understand the possessed data, its potential value and the premise for data utilization need to be in shape before companies can monetize their data. By collaborating with other actors, often current customers and partners, companies can easier create data monetization solutions.

Bibliography

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

Alkkiomäki, V. 2015. Role of service and data reuse in enterprises. *Acta Universitatis Lappeenrantaensis*.

Amit, R. and Zott, C. 2012. Creating Value Through Business Model Innovation. *MIT Sloan Management Review*, 53(53310):41–49.

Angulo, A. 2004. Supply chain information sharing in a vendor managed inventory partnership. *Journal of Business Logistics*, 25(1):101–120.

Baloise Group 2017. Mobile Versicherer "Friday" lanciert. Available from: https://www.baloise.com/de/home/medien/news/2017/mobile-versicherer-friday-lanciert.html (Accessed 23.08.2017).

Barclays 2016. Barclays helps SMEs access 'big data' to aid growth. Available from: http://www.newsroom.barclays.com/r/3336/barclays_helps_smes_access__big_data__to_aid_growth (Accessed 13.06.2017).

Bean, R. 2017. How Companies Say They're Using Big Data. *Harvard Business Review*.

Benzell, S. G., Lagarda, G., and Alstyne, M. W. V. 2016. The Role of APIs in the Economy.

Bilbao-Osorio, B., Dutta, S., and Lanvin, B. 2014. *The Global Information Technology Report 2014: Rewards and Risks of Big Data*. World Economic Forum.

Birkinshaw, J., Brannen, M. Y., and 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.

Blackmer, S. 2015. GDPR : Getting Ready for the New EU General Data Protection Regulation. Available from: http://www.infolawgroup.com/2016/05/articles/gdpr/gdpr-getting-ready-for-the-new-eu-general-data-protection-regulation/ (Accessed 01.09.2017).

Brown, B., Chui, M., and Manyika, J. 2011. Are you ready for the era of 'big data'? *McKinsey Quarterly*, Pp. 24–35.

Buhl, H., Röglinger, M., Moser, F., and Heidemann, J. 2013. Big data.

Chen, H., Chiang, R., and Storey, V. 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*, volume 16.

Cortet, M., Rijks, T., and Nijland, S. 2016. PSD2 : The digital transformation accelerator for banks. *Journal of Payments Strategy & Systems*, 10(1):13–27.

Davenport, T. H. 2015. 5 Essential Principles for Understanding Analytics. *Harvard Business Review*.

Davenport, T. H., Barth, P., and Bean, R. 2012. How 'Big Data' Is Different. *MIT Sloan Management Review*, 54(1):43–46.

Dent, S. 2017. Foursquare can log your shopping trip in increasingly scary detail. Available from: https://www.engadget.com/2017/03/21/foursquare-is-getting-better-at-selling-your-location-data/ (Accessed 09.06.2017).

Dillet, R. 2014. Nest Uses Its Data To Turn Electric Utilities Into Cash Cows. Available from: https://techcrunch.com/2014/04/18/nest-uses-its-data-to-turn-electric-utilities-into-cash-cows/ (Accessed 31.05.2017).

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

European Comission 2015. Directive on Payment Services (PSD) - European Commission. Available from: https://ec.europa.eu/info/law/payment-services-psd-2-directive-eu-2015-2366_en (Accessed 01.09.2017).

European Parliament and of the Council 2016. Regulation (EU) 2016/679. *Official Journal of the European Union*, (L119):88. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32016R0679 (Accessed 01.09.2017).

Evans, N., Fourie, L., and Price, J. 2012. Barriers to the Effective Deployment of Information Assets: The Role of the Executive Manager. *Proceedings of the European Conference on Management, Leadership & Governance*, 7:162–169.

Faria, M., Linden, A., and Laney, D. 2016. Understand the Data Brokerage Market Before Choosing a Provider. *Gartner*, G00294652(2016-01-07):11.

Fisher, C. W. and Kingma, B. R. 2001. Criticality of data quality as exemplified in two disasters. *Information and Management*, 39(2):109–116.

Fred, J. 2017. Data Monetization – How an Organization Can Generate Revenue With Data? *Tampere University of Technology*.

Gioia, D. A., Corley, K. G., and Hamilton, A. L. 2013. Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, 16(1):15–31.

Gore, A., Harmer, P., Pfitzer, M., and Jais, N. 2017. Can Insurance Companies Incentivize Their Customers to Be Healthier? *Harvard Business Review*.

Govindarajan, V. 2016. Stop Saying Big Companies Can't Innovate. *Harvard Business Review*.

Guba, E. and Lincoln, Y. S. 1989. *Fourth Generation Evaluation*.

Haataja, L.-M. 2015. Payment Services Directive II Effects on Business Models and Strategies. *Aalto University*.

Haug, A., Stentoft Arlbjørn, J., Zachariassen, F., and Schlichter, J. 2013. Master data quality barriers: an empirical investigation. *Industrial Management & Data Systems*, 113(2):234–249.

Holt, T. J. and Lampke, E. 2010. Exploring stolen data markets online: Products and market forces. *Criminal Justice Studies*, 23(1):33–50.

Huang, O. and Laney, D. 2014. How Organizations Can Monetize Customer Data. *Gartner*.

IDC 2016. Europe ' s Data Marketplaces – Current Status and Future Perspectives. (June 2016).

Johnson, M. W., Christensen, C. M., and Kagermann, H. 2008. Reinventing your business model. *Harvard Business Review*, 86(12):1–10.

Kastrenakes, J. 2017. Roomba creator says it 'will never sell your data' after talking about selling your data. Available from: https://www.theverge.com/2017/7/28/16055590/roomba-wont-sell-data-irobot (Accessed 07.08.2017).

Kokalitcheva, K. 2016. Allstate Spins Out Startup Arity Focused on Driver Risk Technology. Available from: http://fortune.com/2016/11/10/allstate-arity-startup/ (Accessed 23.08.2017).

Koutroumpis, P. and Leiponen, A. 2013. Understanding the value of (big) data. *2013 IEEE International Conference on Big Data*, Pp. 38–42.

Laney, D., Faria, M., and Duncan, A. D. 2015. Seven Steps to Monetizing Your Information Assets. Technical Report Gartner.

Leidner, D. and Alavi, M. 2001. Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*.

Lewis, A. and McKone, D. 2016. To Get More Value from Your Data, Sell It. *Harvard Business Review*.

Liew, A. 2007. Understanding Data, Information, Knowledge And Their Inter-Relationship. *Journal of Knowledge Management Practice*, 8(2).

Mackenzie, I., Cohn, D., and Gann, D. 2014. The New Patterns of Innovation The New Patterns of Innovation. *Harvard Business Review*, (February).

Mantelero, A. 2013. The EU Proposal for a General Data Protection Regulation and the roots of the right to be forgotten. *Computer Law and Security Review*, 29(3):229–235.

Manyika, J., Chui, M., Brown, B., Bughin, J., and Dobbs, R. 2011. Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*, (June):156.

Mayer-Schönberger, V. and Cukier, K. 2014. Big Data: A Revolution That Will Transform How We Live, Work, and Think. *International Journal of Advertising*, 33(1):181–183.

McAfee, A. and Brynjolfsson, E. 2012. Big Data. The management revolution. *Harvard Business Review*, 90(10):61–68.

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

Merriam-Webster, I. 2017. Monetization — Definition of Monetization by Merriam-Webster. Available from: https://www.merriam-webster.com/dictionary/monetization (Accessed 13.06.2017).

Michelin 2017. Michelin Solutions. Available from: https://michelin-solutions.com/it/ (Accessed 21.08.2017).

Miles, M. B. and Huberman, a. M. 1994. An expanded sourcebook: Qualitative Data Analysis:.

Moore, S. 2015. How to Monetize Your Customer Data - Smarter With Gartner. *Gartner*.

Najjar, M. and Kettinger, W. 2013. Data Monetization: Lessons from a Retailer's Journey. *MIS Quarterly Sloan Management Review*.

OECD 2013. Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data". *OECD Digital Economy Papers*, (222):1–44.

Ohm, P. 2010. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, 57(6):1701–1777.

Opresnik, D. and Taisch, M. 2015. The value of big data in servitization. *International Journal of Production Economics*, 165:174–184.

Osterwalder, A. and Pigneur, Y. 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*.

Osterwalder, A., Pigneur, Y., Bernarda, G., and Smith, A. 2014. *Value proposition design: How to create products and services customers want*.

PatientsLikeMe 2016. How does PatientsLikeMe make money. Available from: https://support.patientslikeme.com/hc/en-us/articles/201245750-How-does-PatientsLikeMe-make-money- (Accessed 10.07.2017).

Rich, R. 2016. Monetizing customers' data. Available from: https://inform.tmforum.org/features-and-analysis/2016/02/8-ways-service-providers-can-monetize-customers-data/ (Accessed 16.07.2017).

Ritchie, J. and Lewis, J. 2014. Qualitative Research Practice: A Guide for Social Science Students and Researchers. *Qualitative Research*, P. 356.

Robson, C. 2007. *Real World Research: a resource for users of social research methods in applied settings.*

Rossman, J. 2016. *The Amazon way on IoT : 10 principles for every leader from the world's leading Internet of things strategies.*

Salmony, M. 2014. Access to accounts: Why banks should embrace an open future. *Journal of Payments Strategy & Systems*, 8(2):157–171.

Saunders, M., Lewis, P., and Thornhill, A. 2009. *Research Methods for Business Students*, volume 5th.

Smith, D. J. 2013. Power-by-the-hour: The role of technology in reshaping business strategy at Rolls-Royce. *Technology Analysis and Strategic Management*, 25(8):987–1007.

Spijker, A. v. 2014. *The New Oil: Using Innovative Business Models to turn Data Into Profit.* Technics Publications.

Stringer, R. 2000. How to Manage Radical Innovation. *California Management Review*, 42(4):70–88.

Strong, D. M., Lee, Y. W., and Wang, R. Y. 1997. Data quality in context. *Communications of the ACM*, 40(5):103–110.

Tamro 2017a. Solutions for pharmacies. Available from: http://www.tamro.fi/en/SolutionsAndProducts/ForPharmacies/Pages/Solutions-for-pharmacies.aspx (Accessed 15.08.2017).

Tamro 2017b. Tamro HANDY. Available from: http://www.tamro.fi/en/SolutionsAndProducts/ForSuppliers/Forpharmaceuticalcompanies/Pages/Tamro-HANDY.aspx (Accessed 15.08.2017).

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

Visconti, R. M., Larocca, A., and Marconi, M. 2017. Big Data-Driven Value Chains and Digital Platforms: From Value Co-Creation to Monetization.

Vives, X. 2008. Innovation and competitive pressure. *Journal of Industrial Economics*, 56(3):419–469.

Weber, R. H. 2010. Internet of Things - New security and privacy challenges. *Computer Law and Security Review*, 26(1):23–30.

Wessel, M. 2012. How Big Companies Should Innovate. *Harvard Business Review*.

Whitmore, T. 2016. Data Monetization: Who's Doing It Right (Now) at a glance Data Monetization. *Blue Hill Research*.

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

Woerner, S. L. and Wixom, B. H. 2015. Big data: Extending the business strategy toolbox.

Wolfe, J. 2017. Roomba vacuum maker iRobot betting big on the 'smart' home. Available from: https://www.reuters.com/article/us-irobot-strategy/roomba-vacuum-maker-irobot-betting-big-on-the-smart-home (Accessed 07.08.2017).

Yin, R. K. 2009. *Case study: design and methods.* Sage Publications.

Ylijoki, O. and Porras, J. 2016. Perspectives to Definition of Big Data: A Mapping Study and Discussion. *Journal of Innovation Management*, 4(1):69–91.

Yousif, M. 2015. The Rise of Data Capital. *IEEE Cloud Computing*, 2(2):4–4.

Zeithaml, V. 1988. Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing*, 52:22.

Zott, C., Amit, R., and Massa, L. 2011. The Business Model: Recent Developments and Future Research. *Journal of Management*, 37(4):1019–1042.

Appendix A

Interview structure

Data monetization interview

Intro 5 min

- Introduction
- The background of interview and the topic
- Is it ok to record?

Data and how it is used currently 20 min

- Your role briefly
- The role of IT? Any analytics team?
- What kind of data is gathered currently?
    - Who owns the data?
- How the gathered data is used currently?
    - To enhance processes and decision making?
    - To improve current products and services?
- How do you use others' data? From who you buy data?
- How valuable is the gathered data?
    - What makes it valuable?
    - Is it valuable also for other stakeholders?

Data monetization and barriers 30 min

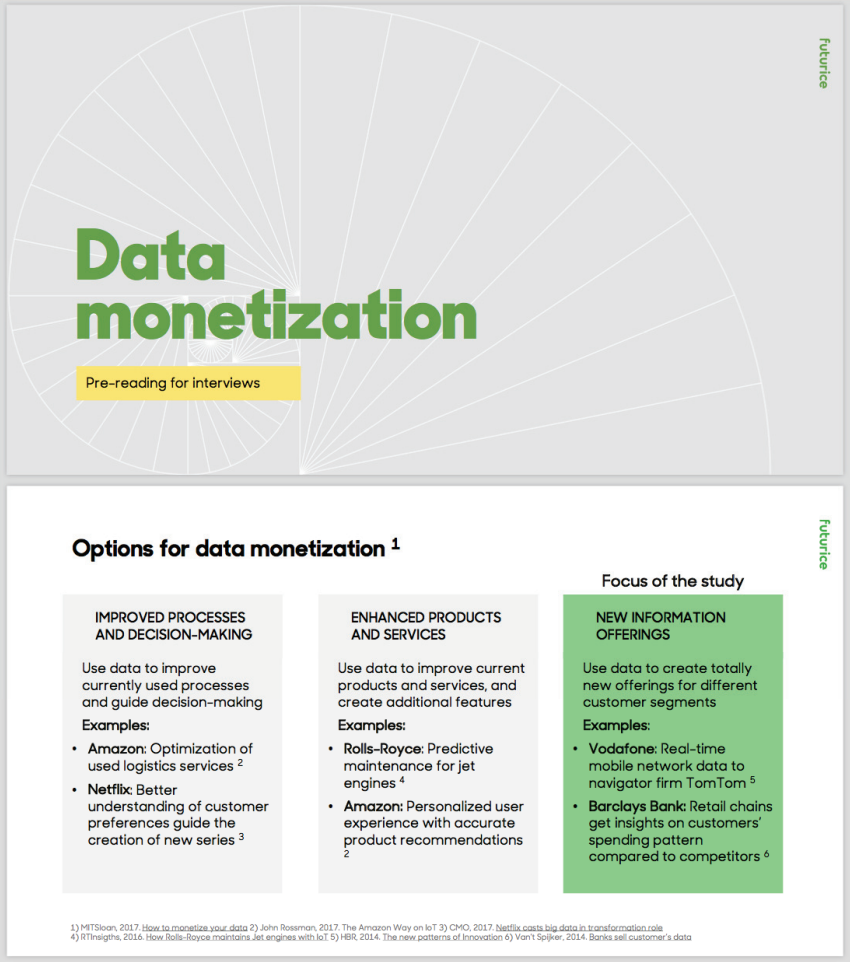
- How do you monetize data currently?
    - Do you sell it to other companies?
    - Do you provide analysis for others?
    - Do you utilize it in some other ways?
- Does somebody else use your data for something?
- Is your data shared in any ways? E.g. Open data projects?
- What should happen, that you could monetize or even utilize your data in new ways? What prevents this kind of action?
- Tell me about your data competences
    - How realistic do you think it is to utilize the data in new ways?
    - How realistic it is to monetize data in new ways?
    - How do you see your data utilization and know-how, compared to:
        - Your own industry?
        - Other publicly listed Finnish companies?
    - What kind of goals do you have for data monetization or utilization? How important it is for you?
    - How do you have prepared yourselves for the future regarding data utilization?

End words 5 min

- Small recap on things said:
    - The current situation of monetization
    - Biggest barriers
- Do you have something you still want to say?
- The next steps on the subject
- Thanks

Appendix B

Pre-reading material for interviews

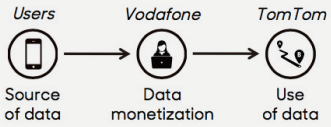


futurice

New information offerings

- Data is gathered in one operational process and monetized by providing value to another customer
- Data is sold directly or it is refined into insights or services
- Data is valuable for other companies due to its historical (insights, predictions) or real-time (data combination) value

Vodafone sells real-time mobile network data to TomTom



futurice

Examples

Company	Case
Adara	Advanced customer segments and profiles from partners' data are used to provide highly-specific services and insights. Partners can be airlines, hotels, and travel agents, who value accurate advertising [1]
Google	Sell insights of Nest smart home usage to utility firms considering energy, appliance and utility aspects of home users [2]
Barclays Bank	Provide retail companies insights about customers' spending pattern in their stores when compared to competitors [3]
Tamro	Service for pharmaceutical suppliers to get information about customers' spending regarding their drugs. In addition provide pharmacies data to compare their sales against their competitors [4]
Foursquare	Service for companies to create highly-targeted ads based on people's location and route history [5]

1) Van't Spijker, 2014. The new oil 2) Techcrunch, 2014. Nest uses its data to turn electric utilities into cash cows 3) Van't Spijker, 2014. Banks sell customer's data 4) Futurice, 2017. Six inspirational ways to make money with data. 5) Wired, 2016. Foursquare's plan to use your data to make money

futurice

Examples

Company	Case
Barclays Bank	Provide SME companies insights and comparison of financial KPIs of other similar businesses in similar locations [1]
Toyota	Sell traffic data from cars to municipalities, corporate delivery fleets and city authors for infrastructure development and route optimization [2]
Quandl	Sell real-time insights to investors about commodities based on marine vessels' movement and harbors' cargo data [3,4]
PatientsLikeMe	Patient-created data considering sickness information and experiences, is used to produce anonymized data for partners, such as pharmaceutical companies and medical device manufacturers, who improve their products according the experiences [5]

1) Finextra,2017. Barclays Bank launches data insights service for SMEs 2) HBR, 2017. To get more value from your data, sell it 3) Quandl, 2017. Monetize alternative data 4) McKinsey, 2013. Are you ready for the era of big data 5) PatientsLikeMe, 2015. How does PatientsLikeMe make money