UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

From Data to Action

Data-Driven Decision Making in Product Development

Mattias Axelsson

Graduate School
Master's Degree Project
Innovation and Industrial Management
Supervisors: Madeleine Englund and Rick Middel

FROM DATA TO ACTION: Data-Driven Decision Making in Product Development
By Mattias Axelsson

School of Business, Economics and Law, University of Gothenburg
Institute of Innovation and Entrepreneurship
Vasagatan 1, SE 405 30, Gothenburg, Sweden

Acknowledgements

I would like to express my appreciation to all who have helped me with this thesis. First, I would like to thank my supervisors Madeleine Englund and Rick Middel for their support and feedback. I would also like to thank my contact persons at Lynk & Co for their support and for giving me this opportunity. Furthermore, I also want to thank everyone who has participated in this study for their time and contributions to this thesis. Finally, I also want to thank the students in my supervision group for their feedback and encouragement.

Abstract

With increasing amounts of data being available, companies are investing to become data-driven with goals of increased performance, productivity, and profit. However, few companies are able to succeed with their investments. This might be due to a disproportionate focus on data and technology compared to the necessary internal processes and culture needed to leverage it in decision making. Thus, the purpose of this study is to explore the extent to which data-driven decision making can be used in product development. Using a single-case study research design with semi-structured interviewing, this thesis explores data-driven decision making as well as the enabling factors of it. Through a thematic analysis, four major themes emerge. These are *the process of data-driven decision making*, *the collective process of data-driven decision making*, *enabling factors*, and *data-driven decision making based on assumptions*. This study answers the research question by stating that data-driven decision making can be used in product development to a certain extent. However, to what extent it can be used appears to depend on how well the process of data-driven decision making gets implemented, if a collective effort is made, and to what extent the enabling factors are present.

Abbreviations

Abbreviations	
AI	Artificial intelligence
DDDM	Data-driven decision making
HiPPO	The highest-paid person’s opinion
KPI(s)	Key performance indicator(s)

Table of Contents

1. Introduction

This chapter introduces the topic of data-driven decision making related to product development. This is followed by a problem discussion which leads to the purpose and research question of the study. Lastly, the research setting, the delimitations, and the disposition of this thesis is presented.

1.1 Background

"*Without data, you're just another person with an opinion*", as said by W. Edwards Deming (European Training Foundation, 2018), reflects a view that data can help separate opinions based on intuition from objective facts. This quote is especially relevant in a world where an estimated 463 exabytes of data will be created every day in 2025 (Desjardins, 2019). This large volume of data is constantly generated from multiple sources and can be accessed in real time (McAfee & Brynjolfsson, 2012). Taken together with advances within business intelligence and analytics, organizations can utilize the massive amounts of data (Chen et al., 2012) which enables them to decide based on data rather than intuition. Using data allows organizations to make better decisions which increases performance (McAfee & Brynjolfsson, 2012) and it is related with increased productivity and profitability (Brynjolfsson et al., 2011). Consequently, companies are investing in data and artificial intelligence with an aim to become data-driven (Bean, 2021).

Being data-driven is a high-level concept encompassing an entire organization (Treder, 2019). It involves managing data sources, classifying data, managing the data quality, and finally, to use the data. In that way, being data-driven involves turning data into information, which leads to insight, which results in value (Treder, 2019). Anderson (2015) elaborates on the final step, creating value, by highlighting the need of acting based on data, i.e., deciding. Thus, an organization is data-driven when its decisions are based on data (Anderson, 2015). Furthermore, these decisions based on the analysis of data are often referred to as data-driven decisions (Provost & Fawcett, 2013) and the need for companies to engage in data-driven decision making is highlighted by several authors (Anderson, 2015; Berntsson Svensson et al., 2019; McAfee & Brynjolfsson, 2012; Nair, 2020; Treder, 2019; Waller, 2020).

One area in particular that could benefit from data-driven decision making is product development within software development. According to Berntsson Svensson et al. (2019), most software is developed using agile methodologies where developments are made in iterations with fast decision making processes. This coupled with the vast amounts of data being generated internally by software-intensive companies puts them in a position where they can engage in data-driven decision making to gain a competitive advantage (Berntsson Svensson et al., 2019).

1.2 Problem Discussion

Despite the investments made by companies (Bean, 2021), few are able to succeed with their attempts of becoming data-driven (Bean & Davenport, 2019). This might be due to a disproportional focus on data and technology compared to the necessary internal processes and culture (Nair, 2020). In fact, being data-driven starts with data and ends with action that generates value (Anderson, 2015). In this sense, data collection, data access, reporting, and analysis are merely pre-requisites. A company that uses data and analysis to generate reports is not data-driven if decisions are still made based on intuition (Anderson, 2015). Conversely, it has been noted that companies often decide based on the opinions of influential people rather than data, which obstructs them from becoming data-driven (McAfee & Brynjolfsson, 2012). Thus, many companies' effort to become data-driven fail due to the managerial challenges associated with implementing data-driven decision making rather than decisions based on intuition (McAfee & Brynjolfsson, 2012; Nair, 2020; Waller, 2020).

Additionally, it has also been argued that data and the associated technology in itself is unable to yield a competitive advantage (Bhansali, 2013; Braganza et al., 2017). Instead, data and technology needs to be combined with, among other, managerial skills and a data-driven culture, where decisions are based on data (Gupta & George, 2016). Moreover, it has also been argued that an increased amount of data, such as big data, may not be that useful if companies are unable to leverage it in their decision making (Ross et al., 2013). Therefore, in order to gain a competitive advantage from data and analytics, data-driven decision making is crucial (Braganza et al., 2017; Gupta & George, 2016; Ross et al., 2013).

Clearly, there are many advantages of data-driven decision making, such as increased performance (McAfee & Brynjolfsson, 2012), productivity and profitability (Brynjolfsson et al., 2011). However, companies appears to struggle in leveraging their data and technology though their decision making processes (McAfee & Brynjolfsson, 2012; Nair, 2020; Waller, 2020) which results in many failed attempts of becoming data-driven (Bean, 2021). Furthermore, this unfulfilled potential of data-driven decision making has also been noted within the field of software product development by Berntsson Svensson et al. (2019). Despite data being available within this field, decisions are often subjective and based on opinions, intuition, and political agendas. Therefore, simply having access to more data does not necessarily result in data-driven decision making (Berntsson Svensson et al., 2019). Thus, there is a practical business problem of establishing data-driven decision making which requires further research.

Simultaneously, there is also a theoretical need for more knowledge regarding data-driven decision making. As noted by several authors, there has been a strong focus on the technical dimensions of working with data but a limited focus on people and culture (Berntsson Svensson et al., 2019; Gupta & George, 2016; Mikalef et al., 2018). As mentioned by Anderson (2015), being data-driven is ultimately about acting on data in order to generate value. Consequently, the limited literature on data-driven decision making from a people perspective needs to be expanded.

1.2.1 Purpose and Research Question

The purpose of this study is to explore how data-driven decision making can be used within product development. More specifically, focus will be on understanding the use of data-driven decision making, and product development will be the research setting. As previously mentioned, there is both a practical, and academic, need of more information regarding the final step of becoming data-driven, namely data-driven decision making (Berntsson Svensson et al., 2019; Gupta & George, 2016; McAfee & Brynjolfsson, 2012; Mikalef et al., 2018; Nair, 2020; Waller, 2020). Based on this, the following research question has been formulated:

To what extent can data-driven decision making be used in product development?

In order to answer this research question, it will be necessary to explore how data-driven decision making is defined in previous literature, gain a better understanding of what the decision-making process looks like, and to explore its enabling factors. This will be done through a qualitative study.

1.3 Research Setting: Lynk & Co

The advantages of data-driven decision making is not only noticed by academia. As mentioned, its promises have resulted in an interest from many companies seeking to improve their decision making. One company with an explicit goal to become data-driven is Lynk & Co. In this section, Lynk & Co will be presented in order to provide a better understanding of the setting in which the research has taken place in.

Formed in 2016, Lynk & Co is an automotive brand that made its debut on the European market in 2020 with a mobility membership business model, and a strong focus on connectivity (Lynk & Co, 2021). Now, as Lynk & Co is transitioning from a start-up, where intuition had a large role in designing their offer, to a scale-up where they have a proven business model and existing customers, they aim to become data-driven.

More specifically, the Business Technology department at Lynk & Co is striving to use data-driven decision making in product development of software intensive products. This department is developing technology and software that helps the rest of the company run its business. Their current product development process follows an agile framework where work items are selected from a backlog to be developed during sprints. Since there might be more work items in the backlog than the scrum team has capacity for during each sprint, work items need to be selected from the backlog. This prioritization of work items is currently being made by the product owners who relies on some data, but mostly experience and intuition. Furthermore, there are also opportunities for data-driven decision making in areas such as deciding on technical solutions, capacity allocation, and general prioritization.

In order to aid decision making in product development, and to ensure that valuable resources are spent on the most important work items, the Business Technology department at Lynk & Co wants to increase their use of data in decision making. Thus, Lynk & Co is an example of a company that has recognized the benefits of data-driven decision making and is in the early stages of implementing it. Apart from being an interesting case to study, they are, similarly to other companies, in need of more knowledge regarding data-driven decision making.

1.4 Delimitations

Although concepts relating to data and analysis will be mentioned, this study does not aim to go into great detail regarding the technical details of it. This means, for example, that specific methods used in analytics will not be discussed, nor will data structures or similar concepts. Instead, data and analysis will be treated as tools in order to understand how it is, and can be used in product development.

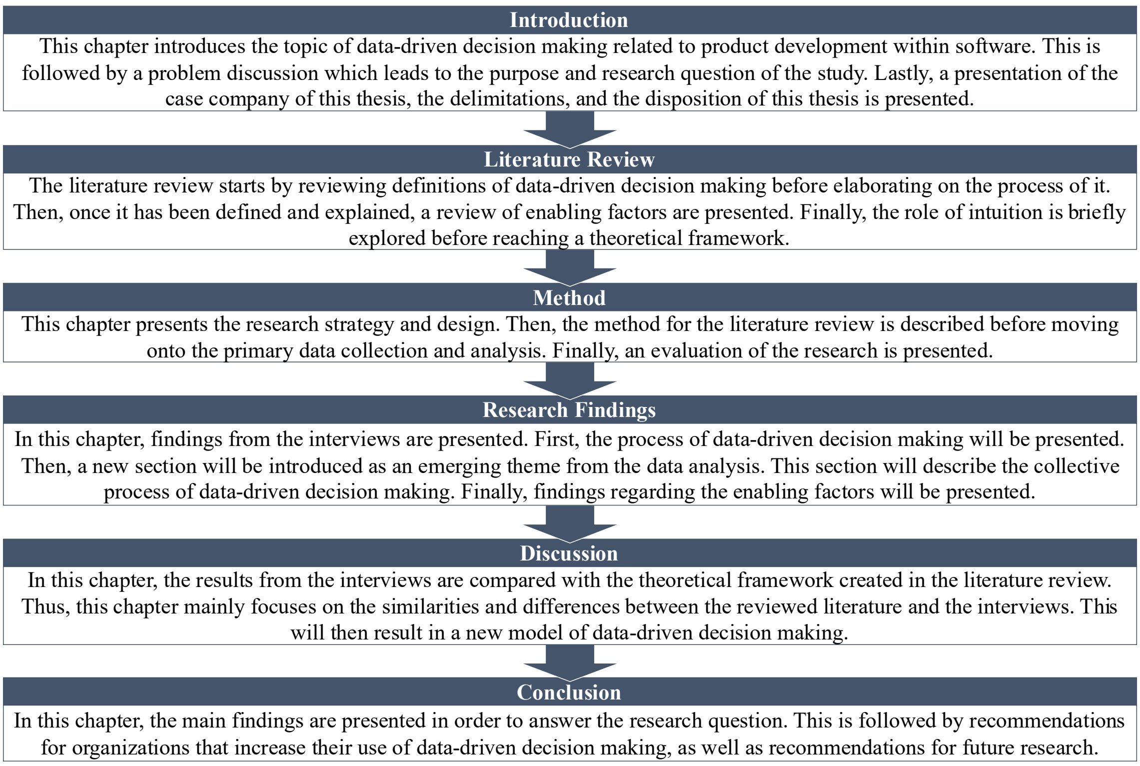
Furthermore, it should be noted that this study will build on interviews regarding data-driven decision making in product development at the Business Technology department at Lynk & Co. Thus, the product development in this thesis will refer to software product development. As a result, attitudes towards data, and the product development itself may be different from other settings. In addition, product development will be treated as a context in which data-driven decision making is taking place, rather than being the focus of this research. Thus, this thesis is delimited to only explore data-driven decision making in the research setting described in section 1.3.

1.5 Disposition

In figure 1, the disposition of this thesis is presented.

Figure 1.

Disposition



2. Literature Review

This literature review starts by reviewing definitions of data-driven decision making (DDDM) before elaborating on the process of DDDM. Then, once it has been defined and explained, a review of factors that might enable DDDM will be presented. Finally, the role of intuition in DDDM will briefly be explored before reaching a theoretical framework of DDDM.

2.1 Defining Data-Driven Decision making

As a starting point, the connection between DDDM and classical decision theory should briefly be clarified. Elgendy et al. (2021) explains this connection by stating that classical decision theory consists of three elements: a decision-making process, a decision maker, and a decision. Within this field, the concept of bounded rationality relates to the belief that humans' cognitive capabilities are incapable of computing the complexities of the world. DDDM adds on to classical decision theory by introducing two further elements: data and analytics. These two elements are added to the three classical elements of decision-theories in order to better understand a complex world and to make more informed decisions.

In addition to explaining the connection to classical decision theory, a definition of DDDM should be clarified. In the literature, DDDM is explained in a number of ways. One area of focus is on requirements that needs to be met. For example, Brynjolfsson and McElheran (2016) require high levels of data availability and usage, the tracking of key performance indicators (KPIs), and long- and short-term targets, to meet their definition of DDDM. More specifically, the targets should be connected to the KPIs and be used to guide action (Brynjolfsson & McElheran, 2016). On the other hand, DDDM can also be defined as a process consisting of multiple steps that start with data and end with a decision (Anderson, 2015; Jia et al., 2015; Mandinach, 2012; Mandinach et al., 2006). In both types of definitions, there is a consensus that DDDM involves basing decisions on data. This is also highlighted in table 1 where multiple definitions are presented.

Table 1

Definitions of DDDM

Definition of DDDM	Author(s) and year
DDDM is an iterative process where data leads to decisions which are implemented and evaluated before starting over.	Mandinach et al. (2006)
Data-driven decisions are based on data rather than intuition.	McAfee and Brynjolfsson (2012)
DDDM refers to a systematic transformation of data to usable knowledge through data collection, analysis, and interpretation, in order to inform decisions.	Mandinach (2012)
DDDM is the practice of deciding based on analysis of data, rather than purely on intuition.	Provost and Fawcett (2013)

DDDM is an iterative process of turning data into knowledge that decisions are based on, while being influenced by an external environment.	Jia et al. (2015)
DDDM requires access to analyses based on trustworthy and relevant data on which decisions are based.	Anderson (2015)
DDDM requires high levels of data availability and usage, the measurement of KPIs, and short- and long-term targets.	Brynjolfsson and McElheran (2016)
DDDM is based on data and analytics combined with the classical elements of decision making; a decision-making process, a decision-maker, and a decision.	Elgendy et al. (2021)

Clearly, DDDM involves decisions based on data. However, a more specific definition is provided by Provost and Fawcett (2013) who defines DDDM as "*the practice of basing decisions on the analysis of data rather than purely on intuition*" (p.53). This definition highlights two key characteristics of DDDM. First, although the strong focus on data in DDDM, data needs to be transformed into knowledge that can be acted on (Anderson, 2015; Jia et al., 2015; Mandinach, 2012; Mandinach et al., 2006; Provost & Fawcett, 2013). This also relates to the view of DDDM as a process. Second, DDDM does not mean that intuition cannot be used, only that the decisions should mainly be based on data (Elgendy et al., 2021; Jia et al., 2015; McAfee & Brynjolfsson, 2012; Provost & Fawcett, 2013). Apart from capturing two key components of DDDM, the definition provided by Provost and Fawcett (2013) also appears to be a common starting point for other articles within the field (Carillo, 2017; Elgendy et al., 2021; Jia et al., 2015). Therefore, their definition of DDDM will be used throughout this paper.

2.2 The Process of Data-Driven Decision Making

As previously mentioned, DDDM is often understood as a process. This process can be described by identifying the steps necessary to take by a decision maker in order to decide. A number of different processes have been proposed in literature. For example, while Chen and Zhang (2014) identify a five step process, Mandinach et al. (2006) propose nine steps, and Jia et al. (2015) propose 11 steps. However, at a high level, all processes can be summarized into 4 steps: data, information, knowledge, and decision. Below, these steps will be described in greater detail.

Data, according to Mandinach et al. (2006), exist in a raw form that does not have meaning in itself. Furthermore, the data can originate from multiple types of sources that are either internal or external to a company (Gupta & George, 2016). Therefore, when a decision needs to be made, it starts with the collection of data (Mandinach et al., 2006). However, it cannot be any data, it has to be the right data. Anderson (2015) explains that the data needs to be trustworthy, unbiased, and timely. Essentially, this means that the data should be relevant for the decision being made, and that it can be trusted by the decision maker and other stakeholders. Moreover, this requires the data to be accessible in the first place (Anderson, 2015).

Once data has been collected, it needs to be organized in a systematic way (Mandinach et al., 2006). In practice, this means that multiple data sources will be arranged or presented in a way that allows further interpretation (Mandinach, 2012). However, at this stage, the data is still merely numbers that have been organized and no interpretation has been done yet (Mandinach, 2012). Therefore, this leads to the next step where data is turned into information.

When data is contextualized, it is turned into information. Thus, it is given meaning that can be used to understand the decision-making situation (Mandinach, 2012; Mandinach et al., 2006). According to Anderson (2015), this can be done through reporting, alerts, and analysis. Reporting refers to extracting data and summarizing it in a report that shows what has happened. Alerting is seen as live reports of what is happening in the present. Analysis goes beyond reporting what has happened by trying to analyze why it has happened in order to make testable predictions. Regardless, it is often recommended that this information is visualized in order to make it more understandable (Chen & Zhang, 2014). Ikemoto and Marsh (2007) also mentions that data can be interpreted both individually and collectively. Here, the collective interpretation of data is seen as more complex, and it results in a mutual interpretation of data based on multiple individuals' expertise. However, information at this stage does not have implications on action (Mandinach et al., 2006).

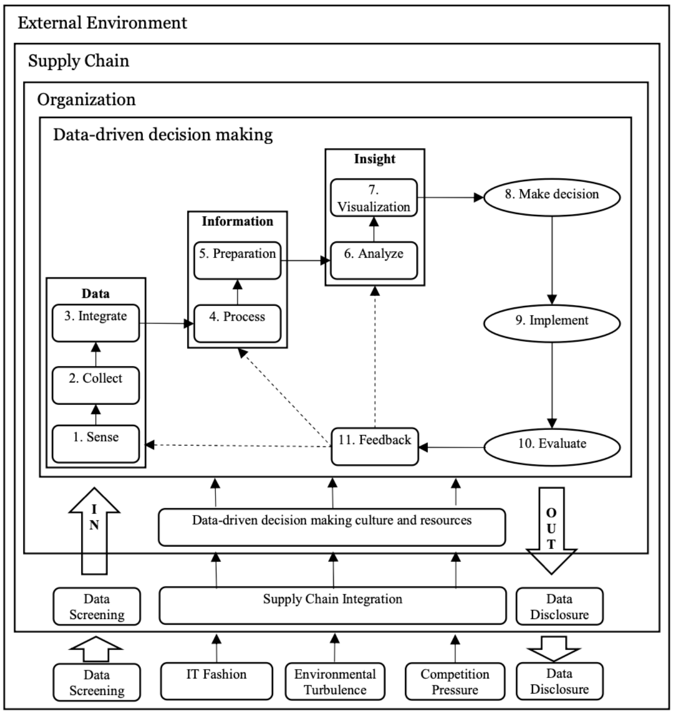
When information is synthesized and prioritized, it turns into knowledge that can guide action (Mandinach et al., 2006). This is an internal process done by the decision-maker where they form a knowledge base (Mandinach, 2012). Or, in other words, the decision-maker reviews the pieces of information and puts them together, decides what information is most useful, and uses that as a basis for a decision. This allows the decision-maker to better understand the decision and what actions can be taken (Mandinach, 2012). According to Anderson (2015), this step is crucial for DDDM since it means that the decision-maker is influenced by the information created from data.

Once the previous steps have been carried out, it results in a decision that gets implemented and evaluated (Mandinach et al., 2006). Thus, a decision in DDDM is an action based on data (Anderson, 2015). However, Ikemoto and Marsh (2007) briefly notes that similar raw data may result in different decisions based on the situation and the decision maker's judgement. Once a decision has been made, Mandinach et al. (2006) explains that the impact of the decision is evaluated which creates feedback loops. This can be related to the specific decision where, for example, additional data may need to be collected (Mandinach et al., 2006), or the process in general (Jia et al., 2015). Thus, DDDM should be seen as an iterative and continuous process (Jia et al., 2015; Mandinach, 2012; Mandinach et al., 2006).

Above, four high level steps of DDDM have been identified: data, information, knowledge, and decision. All of these steps are facilitated by technological tools, data engineering and processing, and data science (Mandinach et al., 2006) and they take place within a business context where the decision making process is affected by the environment it takes place in (Jia et al., 2015). This is clarified in figure 2, which is a model created by Jia et al. (2015) as an adaptation of a similar framework for an educational context, created by Mandinach et al. (2006).

Figure 2

DDDM Process



Note. “Data-Driven Decision Making Process” produced by Jia et al. (2015, p. 6)

Although a decision follows the transformation of data into knowledge, it is important to understand that the process does not describe to what extent the decision is based on the generated knowledge. Mandinach et al. (2006) mentions that it is possible that a decision might not get implemented due to, for example, lack of resources. On the other hand, Jia et al. (2015) includes an external environment in their framework. Though only briefly explained by Jia et al. (2015), this involves organizational culture and resources, which are said to affect the DDDM process. This is aligned with the claim that technology enables analysis, but it is culture that creates a mindset and culture where the findings are noticed, trusted and acted upon (Anderson, 2015). Therefore, in order to understand to what extent DDDM can be used in product development, it important to understand the enabling factors of DDDM.

2.3 Enabling Factors of Data-Driven Decision Making

Above, a process of DDDM was presented and it was noted that the steps alone are not enough to understand to what extent DDDM can be used by organizations. Consequently, there is also a need to review the enabling factors of DDDM. Davenport et al. (2001) had noticed that many companies were gathering data but were not able to make use of it in order to make informed decisions resulting in business value. Therefore, they suggest a model depicting the transformation from data into knowledge and results. This model, shaped as a cone, puts context at the bottom. It is in the context that decisions are made, and therefore it affects the decision making. The context consists of four parts: strategy, skills and experience, organization and culture, technology and data, and organization and culture. Throughout the researched literature, the enabling factors of using data falls into one of these categories, with the exception of basic resources such as time and money brought up by Gupta and George (2016). Therefore, enabling factors of DDDM will be presented according to these five categories.

Furthermore, it should be noted that not all literature is referring to the distinction between the steps of DDDM as described in 2.2. Therefore, the words "data" and "information" may be used interchangeably in coming sections.

2.3.1 Strategy

At its core, DDDM is affected by a firm's strategy, or lack thereof. Without a strategy, companies will not know what they are trying to achieve with their data initiative (Davenport et al., 2001). Instead, a clear vision, goals, and a definition of success is needed in order to guide the use of data in an organization (McAfee & Brynjolfsson, 2012). Moreover, a firm's strategy guides DDDM in two ways.

First, strategy guides data collection. Lin (2018) explains that picking the right metrics is one of the most important steps in DDDM since it mitigates the risk of ending up with too many metrics that do not tell a cohesive story. If data is collected without a clear goal, firms will pick metrics that are interesting on their own but taken together, they do not provide decision makers with the big picture needed to guide action. Gupta and George (2016) further explains that leaders should carefully decide what to measure and which metrics are expected to be used. Additionally, having clear goals to guide data collection also alerts companies when they are missing needed data. Instead of only measuring the data at hand, it is important to understand when new data needs to be collected (Janssen et al., 2017). For example, some goals require new data to be generated through prototyping and experimentation (Martin & Golsby-Smith, 2017). Therefore, the strategic context guides a company in which data to focus on (Davenport et al., 2001).

Second, having a clear strategy helps create organizational support (Davenport et al., 2001) and motivates action (Brynjolfsson & McElheran, 2016). Organizational support is also connected to organization and culture, which will be described later. In terms of motivating action, Brynjolfsson and McElheran (2016) explains that having clear long- and short-term targets gives context to data and motivates action in response to it. In addition, it also allows an organization to monitor their efforts to use DDDM (Jia et al., 2015) Therefore, strategy enables the use of data in the decision making situation.

2.3.2 Skills and Experience

In the process of turning data into knowledge and decisions, skill and experience is needed (Davenport et al., 2001). People skilled in working with data is a crucial asset for companies (McAfee & Brynjolfsson, 2012) and it has been found that more educated workers is positively correlated with DDDM adoption (Brynjolfsson & McElheran, 2016). In fact, Berntsson

Svensson et al. (2019) found that 73% of respondents to their survey indicated that a reason to not use DDDM is due to a lack of understanding of how to use data in a decision. In addition, it has also been found that more experienced decision-makers better understand data which results in better decision quality (Janssen et al., 2017).

More specifically, there are two types of skills that are necessary: technical skills and managerial skills (Gupta & George, 2016). Technical skills refer to the skills needed in order to use technology (Gupta & George, 2016) in the transformation of data into information (Davenport et al., 2001). Although especially important for data scientists, it is also important that people throughout the organization have a basic understanding of the fundamental concepts (Provost & Fawcett, 2013). In fact, technology should be considered everyone's job in order to make use of DDDM (Davenport & Mittal, 2020). This basic knowledge is important when interacting with the produced analyses in order to understand what they convey, but it also allows opportunities to be spotted in order to improve the DDDM process (Provost & Fawcett, 2013).

Managerial skills on the other hand refers to the skill of using the produced information (Gupta & George, 2016). Janssen et al. (2017) explain that this skill improves with experience. When decision-makers start using DDDM, they are likely to experience uncertainty regarding for example how the analysis should be used, and if their decision was correct. However, with more practice of DDDM, the decision-maker will become more experienced. This also relates to the iterative process of DDDM mentioned previously (Jia et al., 2015; Mandinach et al., 2006). Since experience is important, it is suggested that companies should have educational programs in place (Davenport & Mittal, 2020; Waller, 2020). This program should be offered to employees on all levels and departments (Davenport & Mittal, 2020) and preferably in a way that enables employees to use their new skills directly (Waller, 2020).

2.3.3 Technology and Data

Technology and data refer to the underlying hardware and software used when turning data into knowledge, as well as providing end-user access (Davenport et al., 2001). This part of DDDM is one that has gotten plenty of focus in previous literature (Berntsson Svensson et al., 2019), and since this paper is delimited to treat data as a tool, explaining specific hardware and software in great detail is outside the scope of this paper. Instead, it should be noted that it is important that the collected data is trustworthy, timely, and accurate (Anderson, 2015). Therefore, it is important that the knowledge about the data and related analyses is transferred to end users as it helps them understand how the data can be used (Janssen et al., 2017; Power, 2016).

Apart from the quality of data and analyses, Berntsson Svensson et al. (2019) reports that the most common barrier to using data in decision making is that data is not available, or that too much data is available. This relates to two issues that needs to be solved: data access, and data visualizations. First, data access is crucial as little analysis can be done without data (Waller, 2020) and decision-makers are unable to engage in DDDM without data (Anderson, 2015). Second, too much data in a decision making situation can cause confusion and require filtering in order to use data (Berntsson Svensson et al., 2019). This relates back to strategy where it was mentioned that metrics needs to be selected with care in order to collect data that, taken together, helps the decision maker decide (Lin, 2018). Furthermore, visualizations and storytelling is suggested methods of supporting decision-makers as it translates data into a language that all stakeholders can understand (Berntsson Svensson et al., 2019; Davenport & Mittal, 2020).

2.3.4 Organization and Culture

While technology and training can enable analysis, it is the organizational culture that encourages decision makers to notice, trust and use data (Anderson, 2015). This cultural aspect of DDDM is mentioned by several authors, and it is seen as one of the key enablers of DDDM (Davenport & Mittal, 2020; Gupta & George, 2016; McAfee & Brynjolfsson, 2012; Waller, 2020). For example, Carillo (2017) mentions that without a shared vision and culture, the implementation of DDDM will lead to friction between employees. Furthermore, culture is often seen as the factor that is the ultimate obstacle to DDDM (Anderson, 2015; Bean & Davenport, 2019; McAfee & Brynjolfsson, 2012). Below, it will be explained why culture can act as a barrier to DDDM, what a culture that supports DDDM looks like, and how it can be built.

Brynjolfsson and McElheran (2016) conducted a survey on the adoption of DDDM and found that greater tenure of employees negatively correlates with the adoption of DDDM. In addition, they also found that DDDM is less used by CEOs than others in an organization. They hypothesize that these individuals might have high influence within their organizations and are in less need of providing data when motivating a decision. This relates to how a company's culture can act as a barrier to DDDM. McAfee and Brynjolfsson (2012) explain that HiPPO (the highest-paid person's opinion) is a concept used to explain that many decisions are based on an influential person's intuition, rather than on data. This is an issue since companies that rely on the HiPPO rather than data is unlikely to gain any returns on their data initiative (Gupta & George, 2016). However, even when companies are claiming that they are using DDDM, this might not be the case. Instead, McAfee and Brynjolfsson (2012) noticed that many decisions were made based on intuition and later backed up with data that justified the decision. Another cultural barrier to DDDM relates to trust between departments. Janssen et al. (2017) found that data and knowledge sharing is low in some organizations. In order to avoid data silos, there must be a data-sharing culture (Anderson, 2015). Therefore, an organizational culture that enables collaboration and knowledge sharing is crucial for DDDM (Janssen et al., 2017).

Apart from data-sharing, an organizational culture that supports DDDM is characterized by a shared belief that data can be trusted and should be used in decision making (Anderson, 2015). This means that the entire organization needs to value decisions based on data (Davenport et al., 2001), as well as a shifting from basing decisions on intuition to basing them on data (McAfee & Brynjolfsson, 2012). This relates to moving away from the HiPPO and towards decisions based on data (McAfee & Brynjolfsson, 2012) as well as increasing collaboration and data-sharing between departments (Anderson, 2015; Janssen et al., 2017).

In order to establish a data-driven culture, leaders should get into the habit of asking what the data says when making a decision and where the data came from (McAfee & Brynjolfsson, 2012). This helps establishing a data-driven culture since it sets expectation to use data throughout the organization (Davenport & Mittal, 2020; Waller, 2020). Additionally, cross-functional collaboration needs to be maximized to ensure that all decision-makers have access to data (McAfee & Brynjolfsson, 2012). Waller (2020) suggests that boundaries between business units and data scientists should be highly porous. This allows the business units to get closer to data and understand it better. Simultaneously, it also allows the data scientists to better understand the business needs. This is also noted by Janssen et al. (2017) who also mentions that this increases the knowledge of what can be done with data.

2.3.5 Basic Resources

Finally, it should also be noted that companies needs to invest in basic resources such as time and money in their DDDM initiatives (Gupta & George, 2016). In fact, greater investments in IT is positively correlated with DDDM adoption (Brynjolfsson & McElheran, 2016). Adding onto this, it is also possible that a decision that has been reached through DDDM might not get implemented due to a lack of resources (Mandinach et al., 2006). Therefore, a company's ability to invest time and money also shapes its possibility to use DDDM.

2.4 Data-Driven Decision Making Based on Assumptions

Despite the promises of DDDM, it should be noted that it is sometimes criticized. Martin and Golsby-Smith (2017) claims that this scientific approach restricts strategic options and impedes innovation. Furthermore, they argue that it is impossible to make decisions about the future by analyzing historical data. For example, they mention that in order to develop products that change consumers' behavior, imagination and creativity is needed in order to create something new. And for this, there is no historical data. Instead, they suggest that there is a need to create new data and to experiment. In this sense, they argue that access to existing data is less important than the ability to create new data through, for example, prototyping.

However, Davenport (2013) explains that intuition has an important role in data-driven organizations. More specifically, intuition should be used when developing hypothesis as well as when deciding what metrics to track. However, what separates the data-driven from acting solely on intuition is that data is created in order to test hypotheses, and that the effectiveness of selected metrics is measured. Therefore, Davenport (2013) assures that intuition still have a role in DDDM. This is also confirmed in the definition of DDDM provided by Provost and Fawcett (2013) as discussed in 2.1. Furthermore, this interplay between intuition and data-driven analysis might be a key to success (Davenport, 2013). In fact, Lichtenthaler (2018) claims that for the foreseeable future, even the most advanced artificial intelligence (AI) generates the best results in creative work when treated as complementary to the human decision maker. Therefore, Davenport (2013) suggests that an ultimate key to success within data-driven decision making is to find the right mix of intuition and data-driven analysis.

Therefore, it should be noted that there are limitations in using historical data in DDDM (Martin & Golsby-Smith, 2017). Understanding these limits, and knowing when new data needs to be generated, is therefore something that needs to be done by organizations using DDDM (Davenport, 2013; Martin & Golsby-Smith, 2017). However, it does not mean that DDDM has no space for intuition and therefore is unsuitable for creativity (Davenport, 2013).

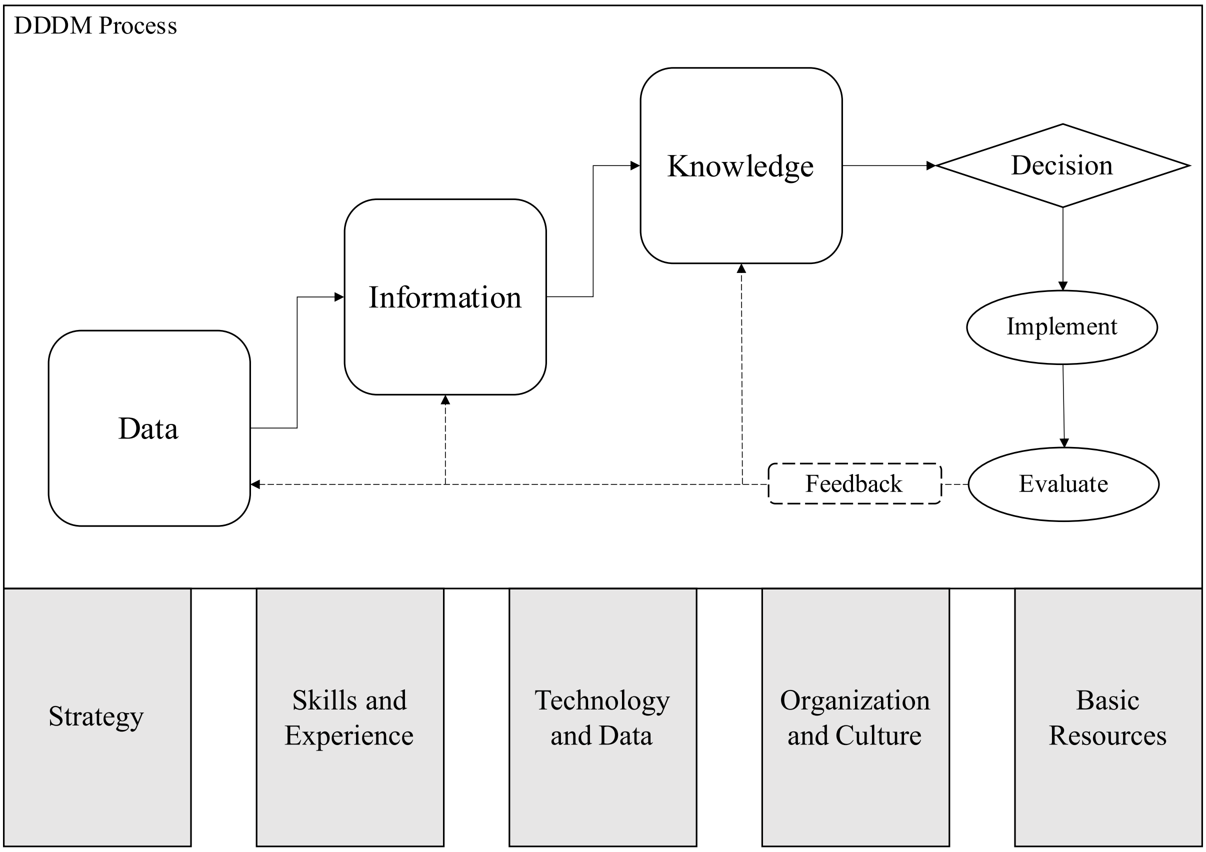
2.5 Key Takeaways

This section started by defining DDDM as "*the practice of basing decision on the analysis of data rather than purely on intuition*" (Provost & Fawcett, 2013, p. 53). Then, the process of DDDM was explained. On a high level, this process starts with raw data which is transformed into information and knowledge that a decision can be based on (Jia et al., 2015; Mandinach, 2012; Mandinach et al., 2006). However, it was noted that the external environment affects the DDDM process (Jia et al., 2015), and that these factors needs to be considered as they affect to what extent DDDM can be used (Davenport et al., 2001). Based on the frameworks created by Davenport et al. (2001) and Gupta and George (2016), five groups of factors were identified: strategy, skills and experience, technology and data, organization and culture, and basic resources. Although data and technology has gotten the most attention in previous literature (Berntsson Svensson et al., 2019), all parts are needed for DDDM. Finally, the role of intuition in DDDM was discussed. Here, it was noted that intuition has a role in data collection in terms of creating new data (Martin & Golsby-Smith, 2017) and picking metrics to measure (Davenport, 2013).

Based on this, a new framework has been created (figure 3). This framework consists of two parts: the DDDM process, and five pillars supporting it. The process is inspired by frameworks by Mandinach et al. (2006) and Jia et al. (2015), and it includes the high-level steps identified above. This process rests on five pillars consisting of the enabling factors of DDDM. Similar to the framework created by Davenport et al. (2001), this is to illustrate that the enabling factors are fundamental to the DDDM process. In appendix 1, the framework is further explained by a summary of the key findings from the literature review. This theoretical framework will serve as a basis for the rest of this paper.

Figure 3

DDDM Process and Enabling Factors



Note. Figure produced by the author

3. Method

This chapter will start by presenting the research strategy and design. Then, the method for the literature review will be presented before moving onto the primary data collection and analysis. Finally, an evaluation of the research will be presented.

3.1 Research Strategy

In order to gain an understanding of to what extent data-driven decision making (DDDM) can be used in product development, a qualitative research strategy was used. As previously noted, there is a lack of research that focuses on people working with DDDM (Berntsson Svensson et al., 2019; Gupta & George, 2016; Mikalef et al., 2018). In other words, more knowledge is needed regarding how people perceive the use of DDDM. Therefore, research that focuses on understanding the attitudes and behavior of people is needed. Since qualitative research emphasizes the understanding of a social world from the participants' perspectives (Bell et al., 2019), it was a suitable research strategy. Through qualitative research, it was possible to gain an understanding of how people perceived DDDM in their daily work. Therefore, the qualitative research strategy allowed for a better understanding of people's views on using DDDM and how that might depend on different contextual factors.

Furthermore, the exploratory nature of qualitative research allowed for the creation of new theory through inductive reasoning (Bell et al., 2019). This was desirable owing to the identified gap in the literature regarding the focus on people. Therefore, this thesis emphasized the generation of new theory, rather than testing existing ones. With that goal in mind, a qualitative research strategy is the best option (Bell et al., 2019).

3.2 Research Design

For this research, a case study design was used. This entails a deep dive into the specific contexts of a single case in order to analyze it intensively. Although the external validity of case studies often is weak (Bell et al., 2019), findings from cases allows for a close understanding of theoretical constructs (Siggelkow, 2007). This results in case studies providing a solid foundation for inductive reasoning, which results in new theory (Bell et al., 2019). Therefore, the use of a case study was aligned with the exploratory nature of this research.

When conducting a case study, it is important that the case is interesting in itself (Bell et al., 2019). On this topic, Flyvbjerg (2010) mentions that in order to maximize the utility of a single case, atypical cases often provides the most information. As mentioned above, the case studied in this research was the company Lynk & Co. This case was expected to offer plenty of information regarding DDDM in software product development, because they have an explicit goal of becoming data-driven. Since this goal was both clearly stated and under implementation, it was believed that it would be an interesting case for the research question. For example, since DDDM was under implementation, the case company could provide timely insights on how this change was perceived. Therefore, Lynk & Co was a suitable case as it was interesting in itself and had the possibility to provide plenty of information.

Furthermore, Lynk & Co had also expressed a need for more knowledge regarding DDDM in product development. Therefore, studying their unique context as a case would provide the company with relevant knowledge, which ensured a practical application of the research.

3.3 Method for Literature Review

Previous knowledge within the field was identified through a literature review. Following the recommendations of Bell et al. (2019) and Snyder (2019), the literature review helped forming a theoretical foundation for the study which was used in order to better understand where this research could add to existing knowledge. Additionally, the theoretical foundation also helped in the creation of an interview guide which was later used during the primary data collection. Below, the procedure of the literature review will be explained.

The literature review was conducted through a semi-systematic, or narrative, review. DDDM is a research field that has been explored from multiple perspectives, some more technical whereas other focuses on decision making behavior. When a research topic has been conceptualized within diverse disciplines, a semi-structured review can be used in order to synthesize the current state of the knowledge within the field (Snyder, 2019). Unlike the systematic review which requires a research question that is clearly defined (Bell et al., 2019), the semi-structured literature review is suitable for research questions that are not narrowly defined (Snyder, 2019). However, inspiration was taken from the systematic review's ability to perform an unbiased search as well as providing transparency of the process. Therefore, search words, inclusion- and exclusion criteria have been used in this process.

The semi-systematic review was facilitated by a search strategy consisting of search terms, inclusion- and exclusion criteria, and databases to be used. As suggested by Snyder (2019), search terms were based on the research question. The search terms consisted of different combinations of the search words in table 2. It should also be noted that the search words were used with variations in hyphens and American/British spelling. The searches, using these search words, were conducted in February and March 2022 on Google Scholar and the Gothenburg University Library.

Table 2

Search Words

Search Words	
• Data driven decision making	• Analytics
• DDDM	• Product development
• DDD	• Management
• Data based decision making	• Capability
• Evidence based decision making	• Business
• Data	• Culture
• Evidence	• Behavior
• Decision making	• Collective

In order to identify relevant articles, inclusion- and exclusion criteria should be used (Snyder, 2019). These criteria, which can be found in table 3, were selected based on a few different reasons. For example, to only include literature that is peer reviewed, has many citations, or published/written by well-known publishers/authors was decided in order to ensure the quality of the reviewed literature. Simultaneously, articles concerning DDDM in niche fields was excluded if the findings were too specific to that particular context. This was to ensure that the reviewed articles would be relevant to the research question.

Table 3

Inclusion- and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
• Literature concerned with the decision-making process of DDDM. • Literature concerned with DDDM in product development. • Literature concerned with DDDM in a general sense. • Literature concerned with barriers to DDDM. • Literature concerned with opportunities of DDDM. • Literature concerned with enabling factors of DDDM. • Literature concerned with enabling factors of using data.	• Literature about DDDM in niche fields that don't apply outside the specific field. • Literature where data-driven approaches have been used in the research rather than being the researched topic. • Literature that was not public or accessible through the Gothenburg University library. • Literature in languages other than English and Swedish. • Non peer reviewed literature, few citations, and published by unknown publishers.

Finally, reviewed articles were also a source of further data collection by looking at citations. Here, both literature that was cited in reviewed literature, as well as literature that was citing the reviewed literature, was reviewed according to the inclusion- and exclusion criteria. Therefore, additional literature that was not found with the initial search words could be discovered.

3.4 Data Collection

In this study, primary data was collected through semi-structured interviews. This step was carried out after the initial literature review. However, due to the inductive nature of this research, this process was somewhat iterative since new knowledge was discovered through the primary data collection. Therefore, as suggested by Bell et al. (2019), the theoretical foundation was not considered to be fixed after the initial literature review. Instead, it remained flexible throughout the primary data collection in order to allow for new concepts to be discovered further.

Primary data was collected through semi-structured interviews. This method of data collection was chosen due to its flexibility where interviewee's point of view can be explored if they bring up unexpected topics (Bell et al., 2019). This was desired due to the exploratory nature of this research where it was expected that new insights would be discovered. Since new insights were expected to be discovered, it would be necessary to ask open ended questions that could be followed-up with probing questions. The possibility to do this is one of the advantages of semi-structured interviewing (Bell et al., 2019). Furthermore, this study has emphasized a need for deeper knowledge regarding the people side of DDDM. Therefore, qualitative interviewing was chosen due to the possibility to understand a social world from the interviewees' perspectives (Bell et al., 2019).

However, it has been argued that semi-structured interviews can be less flexible in dealing with new topics compared to unstructured interviews due to the interviewer trying to ask all interviewees the same questions (Bell et al., 2019). In order to mitigate this risk, open questions were included in the interview guide that would allow the interviewee to make their own connections. Including open questions is also recommended by Bell et al. (2019) to increase flexibility in interviews. For example, asking interviewees what their ideal scenario in working with data, and how that differed from their current situation, proved to be a question that resulted in varying responses from the interviewees. This increased flexibility in the interview since it allowed for new topics to be discovered.

Simultaneously, conducting semi-structured interviews, rather than unstructured interviews, was a decision based on the belief that using a somewhat specific interview guide would be a helpful tool in conducting the interviews. This means that semi-structured interviews are flexible while still providing the interviewer with a tool that helps in covering all desired topics (Bell et al., 2019). Therefore, an interview guide was created based on the literature review. This can be found in appendix 2.

3.4.1 Sampling Interviewees

As this was a case study, sampling was done on two levels. First, a case company was sampled. The sampling of the case company, Lynk & Co, is described in 3.2. Second, interviewees needed to be sampled from Lynk & Co. In sampling interviewees, purposive sampling was used.

More specifically, the purposive sampling was inspired by a combination of theoretical- and snowball sampling as described by Bell et al. (2019). For this research project, it meant that an initial set of 3 participants were strategically selected based on the research question. This meant that people at Lynk & Co that were involved in product development, the work to become data-driven, or both, were sampled. After each of the initially sampled interviewees had been interviewed, a brief analysis was carried out in order to better understand where more information was needed. This part was combined with snowball sampling as the initial interviewees suggested other people that would be relevant for the research question, who in turn recommended additional people. Therefore, this was, similarly to what is described by Bell et al. (2019), an iterative and continuous sampling of interviewees.

This sampling method assured that flexibility could be maintained as the topic evolved, while also allowing for a strategic sampling where relevant individuals could be interviewed (Bell et al., 2019). This flexibility was desired since it was assumed that the research topic would evolve throughout the research when new aspects of DDDM would be discovered. Therefore, being able to sample new participants throughout the research was seen as a better option than sampling all interviewees a priori.

Since the sampling was not pre-determined, there was no goal regarding sample size. Following the concept of theoretical sampling, theoretical saturation was more important than conducting a specific number of interviews (Bell et al., 2019). Since analysis was carried out simultaneously, it became apparent when the number of new codes developed from each interview was decreasing. This meant that further interviews would not have provided new knowledge for the theory-building, thus indicating theoretical saturation (Bell et al., 2019). Consequently, seven interviews were held with people at Lynk & Co. Further details about the interviewees can be found in table 4.

Table 4

List of Interviewees

Interviewee	Date	Position	Location	Duration	Language
Participant 1	2022-03-29	Product owner	Video call	70 min	English
Participant 2	2022-04-05	Product owner	Video call	50 min	Swedish
Participant 3	2022-04-11	Head of XYZ	Video call	50 min	English
Participant 4	2022-04-08	Head of XYZ	Video call	45 min	English
Participant 5	2022-04-14	Product manager	Video call	50 min	English
Participant 6	2022-04-14	Head of XYZ	Video call	50 min	English
Participant 7	2022-04-21	Service manager	Video call	40 min	English

3.4.2 Conducting the Interviews

After potential interviewees had been identified, they were contacted with a short description of the study and asked if they could participate. As suggested by Bell et al. (2019), participants were offered a copy of the interview guide. After the first interview, this guide was updated to include a short description of the differences between data, information, and knowledge. During the first interview, it was noticed that the word “data” can be interpreted as “information”. Therefore, making this distinction from the outset helped guaranteeing that both the interviewee and the interviewer were referring to the same concepts. For the first interview however, this misunderstanding was corrected immediately after being noticed. Therefore, the difference between concepts was also assured during the first interview.

All interviews were held via video calls. This was mainly due to convenience reasons as it facilitated, for example, interviews with participants in multiple locations without having to travel. Furthermore, as all interviewees were rather used to online meetings, technological and behavioral issues connected to online interviewing mentioned by Bell et al. (2019) could be mitigated. Therefore, the interviews were conducted over video calls.

As mentioned above, an interview guide had been prepared for the interviews. This helped create a structure for the interviews where all key areas could be explored. In addition, having pre-prepared questions facilitated the interviews by minimizing leading, or poorly formulated, questions. However, interviews were held rather flexible and additional questions to probe, clarify, or follow-up were also asked. As a result, new topics could be explored in detail, even if this meant that the structure of the interview guide was abandoned at times. Still, with the intentions of using an interview guide as a helping tool rather than a strict script, it was not considered to be an issue. However, it should be noted that asking questions outside the interview guide increases the risk of poorly formulated, or leading, questions. In order to mitigate this risk, this issue was kept in mind when asking unprepared questions. In addition, if it was noticed that a question was misinterpreted, a clarification was given. At the end of each interview, the participants were asked if there was anything they wanted to add. According to Bell et al. (2019), this is good to include since it allows participants to raise any topics that they have not been able to raise during the interview.

Furthermore, with the consent of all interviewees, the interviews were recorded and transcribed in full. This was done since transcriptions facilitate the data analysis and makes the research more transparent (Bell et al., 2019). In order to allow for the continuous analysis needed for the sampling method, all interviews were transcribed immediately after they were held.

3.5 Data Analysis

Thematic analysis was used to analyze the collected primary data. This was a suitable approach since thematic analysis is useful for analyzing interview transcripts, and it is useful for inductive research as it results in themes and theory creation (Bell et al., 2019). Therefore, a thematic analysis was appropriate for the input, and desired output, of the data analysis. Below, the process of the thematic analysis will be further described.

As previously mentioned, the interview transcripts were coded continuously. Using NVivo, transcripts were first coded into data-centered codes. This initial step of coding was to develop codes that stayed close the data by using similar words and phrases as expressed by the interviewees. For example, this meant that for the interview held in Swedish, the initial coding was also done in Swedish. Coding in this descriptive way close to the data minimized any premature, and possibly, biased analysis of the transcripts (Bell et al., 2019). In addition, since NVivo was used, it was also possible to assign multiple codes to the same passages of the transcript.

After the initial coding, NVivo was used to parse through the created codes and group them into concepts. Then, these concepts could be combined into themes. Through this process, data-centered codes transforms into concepts and themes that has a closer connection to existing literature (Bell et al., 2019). By comparing the emerging concepts and themes with existing literature, it was possible to see how they related to each other. As a result, themes that were closely related to existing theories were named accordingly. However, some categories and themes had no clear connection to the reviewed literature. Thus, these were considered as emerging theories. In appendix 3, a visualization of the thematic analysis can be found.

During the process of coding, extra attention was given to passages of the transcripts that indicated connections between concepts. This was possible by identifying how participants expressed causal connections by using words such as “because” (Bell et al., 2019). By doing this, it was possible to better understand how interviewees perceived the use of DDDM in connection to the developed themes, as well as the connection between themes. As a result, it was possible to analyze how different factors affected the perceived use of DDDM in product development.

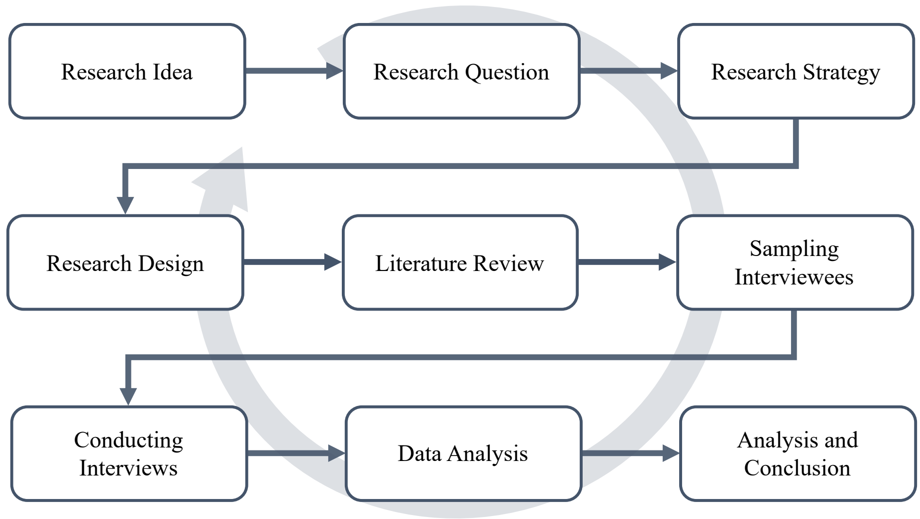
However, Goffin et al. (2019) suggests that data should be coded independently by multiple researchers in order to increase the rigor of the research. Since this study was conducted by a single researcher, that was unfortunately not possible. In order to mitigate this risk, it was decided to use NVivo. As mentioned above, this allowed for the creation of a great number of individual data-centered codes and the same text in transcripts could be coded multiple times. Taken together, this allowed the coding process to accurately transform interview transcripts into codes while minimizing any premature or biased analysis of the transcripts.

3.6 Research Process

The research method described above can be summarized according to figure 4. This figure shows a flow from an initial research idea to a final conclusion based on research findings. However, this flow should also be seen as an iterative process. More specifically, the flow from research question to literature review created a loop. This was due to more knowledge about the field being uncovered through the literature review. Thus, the research question could be more clearly defined. Another loop in this process was from literature review to data analysis. As interviews were conducted and analyzed, the literature review was iterated when unanticipated themes and categories were discovered. A final loop in the process was connected to the sampling of interviewees. As described above, the data analysis determined the need for continued sampling of interviewees. Thus, the sampling of interviewees to data analysis was also an iterative process.

Figure 4

Research Process



3.7 Research Quality

In this section, the quality of this research will be discussed. Since the criteria reliability and validity are better suited for quantitative studies, this discussion will be based on the alternative criteria for qualitative research presented in Bell et al. (2019). These are credibility, transferability, dependability, and confirmability.

3.7.1 Credibility

Credibility refers to how believable the findings are, and it depends on whether good research practice has been carried out, and if it has been confirmed that the social world has been correctly understood (Bell et al., 2019). Since this study has been carried out by following the recommendations by Bell et al. (2019) and the supervisors to this thesis, attempts have been made to follow good research practice as described above. In terms of confirming that the social world has been correctly understood, respondent validation has been used. According to Bell et al. (2019), respondent validation refers to a process to ensure that the findings corresponds to the experiences and perspectives of the interviewees. This was achieved by emailing each participant a copy of the findings from their interview. As a result, participants could verify that the findings corresponded to what they had wished to express. In the case that a participant would not have agreed with a finding, this would have been investigated further in order to understand the reason for it.

3.7.2 Transferability

Transferability relates to external validity, and thus how well the findings can be applied in other contexts (Bell et al., 2019). As noted above, the external validity of case studies is often weak (Bell et al., 2019) and instead its strength lies in the ability to get close to theory and understand causal forces (Siggelkow, 2007). However, Bell et al. (2019) mentions that qualitative research should include extensive descriptions of the research in order to allow others to make their own judgements regarding how well the findings can be transferred to other settings. Therefore, this thesis has aimed at providing a detailed description of the case company, the interviewees, the interview guide, and the research in general.

3.7.3 Dependability

Dependability relates to the trustworthiness of the research, and it can be increased by providing extensive information about the research process (Bell et al., 2019). In the sections above, it has been attempted to provide as detailed information as possible regarding the research process and the choices that has been made in order to increase dependability. For example, by including a visualization of the thematic analysis in appendix 3, detailed and transparent information about the data analysis has been provided. In addition, Bell et al. (2019) also mentions that all records should be kept for auditing, which increases dependability. For this research, all material, such as interview transcripts, has been saved. However, this study has not been audited in the way where others have reviewed this material. Instead, the thesis' supervisors and other students have been involved in the research process by providing feedback on the thesis on a regular basis.

3.7.4 Confirmability

Confirmability refers to the extent to which the research process has not been affected by values or previous beliefs of the researcher (Bell et al., 2019). As mentioned in the sections above, confirmability has been strived for in a number of ways. First, an interview guide was prepared with open questions based on previous literature. Thus, questions could remain free of personal bias, and they were based on previous literature rather than personal beliefs. Second, the initial coding of transcripts was done using data-centered codes which should minimize bias in the coding process.

4. Research Findings

In this chapter, findings from the interviews will be presented. The structure will be similar to the one presented in the literature review. First, the process of data-driven decision making (DDDM) will be presented. Then, a new section will be introduced as an emerging theme from the data analysis. This section will describe the collective process of data-driven decision making. Finally, findings regarding the enabling factors will be presented before presenting participants' views on the limitations of historical data when developing new products.

The thematic analysis of the interviews resulted in the identification of four main themes and a number of categories related to data-driven decision making in product development. As previously mentioned, some of these themes had close connections to existing literature and were therefore named accordingly. However, a new major theme emerged through the thematic analysis. This theme will be called the collective process of data-driven decision making. In appendix 3, the thematic framework developed during the thematic analysis is presented.

In the coming sections, these themes will be described in greater detail. The first section will focus on the process of DDDM by presenting how the interviewees perceived the use of data in decision making on a general level. Then, it will be explained how DDDM becomes a collective process when it is taking place in a setting where multiple people come together in making decisions. The next theme to be presented relates to the enabling factors of DDDM. Finally, the last theme to be presented relates to participants' views on the limitations of historical data when developing new products

Furthermore, it should be noted that all interviews were held within the research setting as described in 1.3. Therefore, the coming sections will present the interviewees' perceptions of DDDM within the context of product development of software-intensive products.

4.1 The Process of Data-Driven Decision Making

Overall, the participants shared an understanding that raw data needs to be collected, turned into information, and then analyzed to make sense of it. For example, participant 1 said that it is not possible to decide based on raw data without first understanding what it means. In line with this, Participant 4 described DDDM as "taking the raw data into information, and then further analyze it to make sense of it". At the same time, it was also recognized that this is an iterative process that includes subjective choices. Below, this process will be explored in greater detail.

According to the interviewees, data can be collected from multiple sources. In terms of external sources, participants 5 and 7 explained that external data can be collected for decision making, but it may be less accessible and useful compared to internal data. For example, participant 7 mentioned that external data could be studies about customer behavior in the US or globally. Since Lynk & Co focuses on the European market, participant 7 was not sure how those studies could be applied. Regarding internal data, participant 3 explained that they have access to the systems and can look directly in the databases. On this topic, participant 1 explained that they have plenty of data within these systems, but they need to collect and process it in order to show it in a user-friendly way.

Regarding the next step of the DDDM process, discussions among the interviewees mainly focused on the challenges of turning data into information that can be understood throughout the company. Although participants mentioned that they are gathering information on their own, it was raised that the process of DDDM is not done by a single individual in isolation. Instead, multiple people come together and play different parts in the process. This topic turned out to be rather important to DDDM and will therefore be treated as a separate theme in 4.2.

In terms of turning information into knowledge and making decisions, the interviewees discussed how the same information gets interpreted in multiple ways depending on the individual. The reason for this, as explained by participant 1, is that there will always be some factors in a decision that are be subjective. As an example, participant 1 mentioned customer experience: “does it offer a good customer experience? What is a good customer experience for me? Probably it is not for you, and the other way around”.

As a result, participants believed that decisions cannot be based completely on data, and that there is room for other factors in the decision making. Although these factors were referred to as intuition at times, several participants also explained that it consists of experience, preference, existing knowledge, and values. Since these factors may differ between individuals, it was noted that similar information may result in different decisions. For instance, participant 6 explained how values between departments results in different opinions based on the same information.

> *We place a different priority on things than they might. They really value the UI experience, and we value functionality working. So, there are sometimes when we disagree on that.*
> Participant 6

4.2 The Collective Process of Data-Driven Decision Making

Based on the interviews, it was clear that many decisions within product development at Lynk & Co are not taken in isolation by a single individual. Some decisions are taken together, some decisions are influenced by stakeholders, and some decisions will affect stakeholders. As a result, participants 1, 2, 3, and 4 discussed the need of information to be clearly defined throughout the organization. These discussions resulted in this emerging theme consisting of two categories, the need for clear and shared definitions, and how they are created.

4.2.1 The Need for Clear and Shared Definitions

According to the interviewees, a crucial step in DDDM is to ensure that there is a shared business language throughout the organization. As mentioned above, decisions are not made in isolation, and a common business language is needed to ensure efficient discussions. As expressed by participant 3, “If we don't speak the same language, we cannot have a discussion.

We can still say things to each other, but we will get nowhere”. This refers to a need of clear and shared definitions, and the interviewees recognized this as a fundamental part of DDDM.

> *You need to make sure that you speak the same language throughout the organization. That we have the same definitions. That we know what we mean when we say something and that we don’t just make assumptions.*
> Participant 2

From the interviews, it was clear that definitions refer to an agreement of what, for example, a KPI means and how it is measured. According to participant 2, this relates to understanding information. Although reports might be available, it might be difficult to understand the information if definitions are not clear. In fact, several interviewees mentioned how common terms used in the daily work are poorly defined. Although it is possible to have a high-level discussion on a topic, it gets difficult to define it in a KPI.

For instance, participant 3 described how a metric such as newsletter registrations can have two different values depending on how it is measured. One option could be to use Google Analytics data to count how many times people have signed up for newsletters on the website. Another option would be to use customer relationship management (CRM) data to see how many email addresses allows newsletters. Due to a number of reasons, these values might be different. Therefore, without a clear definition of what is meant by newsletter registration, two interpretations can be made. Furthermore, participant 3 also mentioned that this issue gets even more complicated for more complex processes. Therefore, clear, and shared definitions were recognized as important in order to avoid confusion regarding reports and other pieces of information.

In addition, as explained by participant 4, a lack of definitions results in assumptions that are not necessarily shared across the company. This, in turn, affects the usefulness of information since it is not clear what it means. Participant 4 also mentioned that this may result in multiple reports on the same topic but with different underlying assumptions. Then it is difficult to know which report is the correct one. Furthermore, a lack of clear definitions may also result in decreased trust in data and tension between departments. This is explained by participant 3:

> *If people haven't agreed yet what, for example, membership should mean. If we start sharing how many members we have, of course that creates tension within the company because people would have different expectations. And they start questioning those numbers.*
> Participant 3

4.2.2 Creating Clear and Shared Definitions

Achieving clear definitions is seen as a result of understanding business processes, and shared definitions are seen as a result of discussions. According to the participants 2, 3 and 4, understanding processes refers to being able to pinpoint where in a process a KPI should be measured. Similar to the newsletter registration example above, participant 2 mentioned an example related to orders. In both of these examples, participants 2 and 3 explained that it is important to define at what point in a process data should be collected, and from which source. However, as mentioned by participant 4, understanding processes and thus being able to create definitions is a task that varies in complexity. While some definitions are rather intuitive, others are more complex. As a result, getting clear definitions is easier in some areas than in others.

> *It's easier in a way to relate to and find direct connections in certain areas. Like within financial performance, or counting time, or how much it takes to do something, or how much usage there is on a web page, or whatever. It's easier to come to those figures and become data-driven and take actions on than in certain areas where it's harder to come to the set of figures.*
> Participant 4

Achieving shared definitions, on the other hand, is done through discussions. Participant 3 explained that it is important that stakeholders who will use the information should participate in creating definitions. This relates to the common business language which needs to be shared across an organization. However, participant 3 also mentioned that there is a risk of lengthy discussions when creating definitions. Therefore, participant 3 stated that these discussions must result in an agreement, so that information about the definitions can be shared throughout the company.

> *What I feel also, is that sometimes we come to agreements and very soon after, they were questioned again. And then we kind of lost those definitions very early in the process. [...] We will need to come to an agreement [...] and then spread the word.*
> Participant 3.

In addition, participant 3 also explained that they would like to be more transparent in their reporting of information. It was suggested that reports should be publicly available within the company and that they would be as transparent as possible. For example, participant 3 mentioned that the reports should include an explanation of what is measured and how. Similarly, participant 4 had noted that reports which include an explanation of how data was turned into information are easier to use in decision making. Therefore, definitions could be shared together with reports to increase transparency and possibly increase their use in decision making.

4.3 Enabling Factors of Data-Driven Decision Making

4.3.1 Strategy

According to the interviewees, being data-driven is being perceived as a strategic goal. However, participant 2 explains that this goal is rather an ambition than having a clearly formulated end goal. As explained by participant 4, there are several aspects of the goal, but it is ultimately to run the business on figures rather than gut feeling, which is intended to result in better decisions. Since this goal is broadly defined, it allows for different interpretations which was clear among the participants when they described how they use DDDM. Participants mentioned different reasons for using data such as helping to guide development (participants 1, 3, 5 and 7), learning about customers (participants 5 and 7), and monitoring (participants 1 and 3). As an example of how the ambition of DDDM gets interpreted, participant 6 explained that data is collected based on the goal of driving conversion on the website. Therefore, participant 6 described how they are mostly interested in data from the offer pages and check-out pages.

> *I think the goal is definitely to drive better conversion. [...] So, I think the majority of the time is spent with check-out data. […] I would say that our primary view is the offer pages to check-out. Because those are very important.*
> Participant 6

Although the goal of DDDM is recognized as a broad goal which allows for many reasons to use data, participant 4 also explained that the end goal of DDDM is related to the broader strategy of the firm. Here, participant 4 mentioned sustainability, experiences, and, ultimately, profitable business. Thus, DDDM was described by participant 4 as "a tool for us to make better decisions and thus run the business better".

4.3.2 Skills and Experience

In terms of skills and experience, discussions among participants mainly focused on the managerial skills of DDDM. Although technical skills were also discussed, these discussions gravitated towards both the collective process of DDDM, and resources as an enabling factor. Both participant 1 and 6 expressed that although they feel comfortable parsing through data to understand it, they are sometimes lacking the resources to actually do it. In addition, participant 3 expressed that the biggest obstacles to understanding information is the lack of definitions.

However, participant 1 did mention that not everyone has the technical skills needed to turn data into information. Still, this was not seen as an issue, but something good. Participant 1 explained that people without this skill can still understand information, and that their understanding will be from a different perspective than the ones who produced the information. A more technically oriented person might understand KPIs from a technical point of view whereas a business-oriented person might understand the same KPI from a business point of view. Therefore, participant 1 did not believe that everyone needs to have the technical skills of transforming data into information.

Regarding managerial skills, all participants expressed eagerness towards using more data in their decision making. Several interviewees mentioned that having data to back up their decisions would make their decisions more transparent, which would makes them feel more comfortable. For example, participant 1 stated that "data-driven decisions make me feel more comfortable and confident with the decisions I take […] because it makes the process more transparent and easier for everyone to understand".

On the other hand, participant 4 believed that data in decision making will not make them more confident in a single decision. Instead, the increased confidence comes from multiple decisions based on data, which makes it possible to adjust and follow up on wrong decisions. At the same time, participant 4 also stated that the goal should not be to eliminate wrong decisions, since wrong decisions can be the result of trying “crazy things”. Then, data helps to better understand those decisions and make better ones in the future.

> *It doesn't make me more confident on one individual decision. But in the total, it makes me more confident. If we make a wrong decision once, we can follow it up and adjust to the right direction. It also makes me more calm that: yes, we will for sure make wrong decisions. That's always going to be the case. If we didn't, then we didn't try crazy enough things. But by having this follow up, then you can change those crazy decisions to better ones.*
> Participant 4

4.3.3 Technology and Data

Regarding technology and data, discussions among the interviewees mainly focused on data access. Here, participants described both situations where they have access to data, and situations where they do not have access to data. For example, participant 6 described how they are using tools to gather and present data regarding the website. This information is made available during meetings, but the team members also have access to the tools themselves if they want to access the data on their own.

On the other hand, an emerging category from the interviews with participants 5 and 7 related to dependencies on other companies which may restrict data access. Since not all product development is made internally in Lynk & Co, some systems that generate data are owned by other companies. As mentioned by participant 5, dependencies on other companies are very common within the automotive industry. Consequently, access to this data is limited to what the other company is willing, or capable of sharing.

> *But we don't own the data, we don't see anything. So, we are reliant on our supplier or partner to provide us with what they are willing to provide. […] I am pretty sure that they can see much, much, much more, but they are not obliged in any way to share this, and are also not willing to share it.*
> Participant 5

This means that there are limitations connected to suppliers. Participant 7 explained that although they are getting access to some data, their supplier is not providing data at a detailed enough level. As a result, participant 7 explained that they can only get an indication of how customers are using the service, rather than getting all data they would like to have. Although participant 5 and 7 get some data, they expressed that they are not getting enough. Therefore, they explained that requirements regarding data access should be defined clearly from the beginning to ensure access. This relates to getting data access from suppliers, but also to make sure that products that are co-developed with partners are required to include data collection from the start.

4.3.4 Organization and Culture

Participants 3, 5, and 6 described that the goal of using data in decision making has resulted in a cultural shift. Together, they described that during the start-up phase of Lynk & Co, they had little data to rely on. Instead, instinct, intuition, and previous experience was more important in the decision making. Here, participant 5 made a general comment that “when you have a startup, you need […] to be very sure of yourself. Sometimes more than what can be proven by hard data”. Now, on the other hand, Lynk & Co have access to data and participant 3 described that they should increase the use of it.

> *We came from a point where we were relying on our instincts, on our previous experience, to take decisions. Because we were about to launch something new, a new business model, a new company. […] But now when we go live and we have one year of data that can support us, we are all very keen to know if our previous experience with our instincts helped us going in the right direction.*
> Participant 3

Although all interviewees expressed positive attitudes towards increased use of data in decision making, participant 6 described that this cultural shift is not happening at the same rate throughout the company. From participant 6, it could be understood that the Business Technology department might be more data oriented than the rest of the company. This difference in decision making causes friction since product teams and stakeholders arrives at different priorities. Therefore, participant 6 believed that this is a cultural shift that needs to happen throughout the company.

> *I think as a company, we are geared to our intuition. But as product teams we are more willing to work with data and use that.*
> Participant 6

However, participant 3 believed that most people at Lynk & Co wants to make the shift, but the lack of clear and shared definitions causes uncertainty which can negatively affect the attitudes towards using data. Therefore, participant 3 saw definitions as affecting the cultural aspect of DDDM. Furthermore, all participants expressed positive attitudes towards DDDM, and they also mentioned that people throughout the company were requesting more data as well. For example, participant 1 explained that stakeholders appreciate, and requests, the transparency that data adds to decisions. Likewise, participant 4 explained that data is being requested during meetings as a basis for decisions.

In terms of a data sharing culture, all participant described a culture that promotes sharing of knowledge. Participant 6 stated that sharing is often done ad hoc by asking other people for relevant insights, and participant 4 mentioned reports that reach a wide audience within the company. However, participants 1, 4, 6, and 7 also mentioned that everything does not need to be shared. They explained that some data is too specific for their function to make sense in other areas of the company. As stated by participant 1, “one function’s report doesn't make sense to the next function’s report”. Likewise, participant 1 explained that they need very specific information about their system, and that data from other systems is not always relevant. On the same topic, participant 7 explained that conclusions from data might be more interesting to share than the data itself. Therefore, participants expressed positive attitudes towards sharing knowledge and data, but they also recognized that data needs to be made understandable for other functions to use it.

> *I don't see that we wouldn't want to share, and that there would be anyone who don't want to share what they're doing. It's more of to making it relevant for others and having time and resources to do it.*
> Participant 4

4.3.5 Basic Resources

Participants 1, 2, 3, 4, and 6 recognized time as important in creating definitions, turning data to information, and to read information. Participant 3 explained that although plenty of time has been invested in creating clear and shared definitions, more time is needed to discuss definitions. Participant 4 described the need for time related to both definitions and to turn data into information in the following way:

> *Time available to sit down, understand and get enough knowledge of a certain area to be able to come from the raw data to real information and real knowledge. And be able to do it focused is, I think, the resources you need to have to make it successfully.*
> Participant 4

Simultaneously, participant 4 explained that time is needed to read and understand what the reports are saying. Something that they are sometimes too busy to do. Similarly, participant 6 explained that they are too busy with other tasks to properly make use of data and give it the necessary attention. Instead, participant 6 suggested that they would need a dedicated expert who could focus on data and bring information to the product owners.

This relates to another resource, people. Participant 6 expressed a need for more people to properly make use of data. A similar need had been observed by participant 1 who decided to bring in two additional people to build a way to track KPIs. However, participant 2 explained that there might be a decreasing marginal effect of involving more people. Thus participant 2 identified a contradiction between involving more people to increase resources, and that more people may result in a less robust solution.

> *Of course, there could be 10 people, but then it is not certain that we can establish a sustainable and robust solution. It could go too quickly, and it makes it difficult to coordinate between people.*
> Participant 2

4.4 Data-Driven Decision Making Based on Assumptions

Finally, the participants also discussed their views on developing new products with data-driven decision making. When developing something new participants recognized that there is no relevant historical data. Instead, participants 3 and 4 explained that data-driven decision making in this context is about testing assumptions. An initial decision might be made based on assumptions, but this should immediately be followed up by data collection. Participants 3 and 4 explain that they would like to collect data from the start in order to follow up on their assumptions. Participant 3 refers to an agile way of developing where they would develop a small part, get data, and understand what the next step should be. Overall, they highlighted that this is an iterative process where they are collecting data in order to learn more and make better decisions in the future.

> *But then I think, making something for the first time, we can make a decision that: OK we make this new service, or we make this new function based on the assumption that we think is right. That's just fine in my opinion. The goal of it should be to follow it up. And after a while, look back to the assumptions and see whether they were right or wrong. Then you can learn out of it, build some new knowledge, and make a better decision next time.*
> Participant 4

5. Discussion

In this chapter, the results from the interviews will be compared with the theoretical framework created in the literature review. Thus, this chapter will mainly focus on the similarities and differences between the reviewed literature and the interviewees' perceptions of DDDM. This will then result in a new model of DDDM.

5.1 The Process of Data-Driven Decision Making

From the literature review, it was clear that DDDM is an iterative process that starts with data and ends with a decision (Jia et al., 2015; Mandinach, 2012; Mandinach et al., 2006). Similarly, participants expressed that raw data needs to be transformed into information that can be used to inform decisions. Furthermore, Mandinach et al. (2006) stated that the impact of a decision results in feedback loops which affects future data collection and decisions. Although this area was mostly brought up under 4.4, participants mentioned that there is a learning connected to DDDM which is believed to improve future decisions. Thus, at a macro level, it appears that the literature and the participants' views on the DDDM process are aligned. However, to better understand the process of DDDM, this will be investigated more in-depth.

Starting with data, it was noted that data exists in a raw state that has little meaning in itself (Mandinach et al., 2006) and that it can be collected from both internal and external sources (Gupta & George, 2016). While it was not surprising that participants 1, 3, and 4 recognized that data exists in a raw state, it was interesting that participants 5 and 7 explained that external data might not be so accessible and useful. According to Anderson (2015), it is important that the collected data is relevant, accessible, and trustworthy for the decision making situation. However, based on the participants' view on external data, it could be argued that it does not fulfill these requirements. For example, participant 7 described that that they are sometimes lacking studies on the European market and have to resort to studies conducted on a global level. This suggests that the data may not be completely relevant. Conversely, participants 5 and 7 explained that data owned by suppliers may result in data being inaccessible. Despite external data potentially being less accessible and relevant, it should still be noted that it might have been the best option, since the other option could be to have no data. Therefore, it might be that Anderson (2015) describes an optimal situation, but that DDDM can still be possible as long as the decision makers are aware of the limitations of the data.

As mentioned in 4.2, turning data into information was discovered to be part of the major theme about the collective process of DDDM. Therefore, this step will be analyzed in the section below.

Turning information into knowledge that guides a decision was explained as an internal process where the decision maker makes sense of information (Mandinach, 2012; Mandinach et al., 2006). However, the literature offers little explanation on how this internal process is carried out. Based on the interviews, it was discovered that the same information can be understood in a number of ways, which results in different decisions. According to the definition of DDDM as "*the practice of basing decisions on the analysis of data rather than purely on intuition*"

(Provost & Fawcett, 2013, p. 53), it is clear that some level of intuition is used in DDDM. This intuition is probably what results in the different understandings of the same information. However, based on the participants' discussions, intuition in this sense could be understood as a combination of experience, preference, existing knowledge, and values. Therefore, at the final stage of the DDDM process, information gets combined with a person's own understanding and values to create knowledge.

This contributes to the literature by identifying a missing piece to the DDDM process. The processes described in the reviewed literature (Jia et al., 2015; Mandinach, 2012; Mandinach et al., 2006) shows the DDDM process from the perspective of data, rather than showing how different components feed into a decision. Therefore, in order to better show that there is likely to be some subjectivity in the knowledge creation and decision, it should be added to the model. Furthermore, Ikemoto and Marsh (2007) noted that similar raw data may result in different decisions based on the decision maker's judgement. This is important since, as noted by participant 6, people are likely to have different experiences and values. Thus, it is likely that the same information will result in multiple understandings, which might cause disagreements. Therefore, adding an intuition component to the model highlights that the process will not always result in the same decision, even if it has the same data as a starting point. Furthermore, when decisions are made by multiple people, it is likely that they bring different values, experience, and previous knowledge into the decision-making situation. Therefore, this component could also highlight that a decision might, to some extent, also be the result of an agreement. A suggestion of what this could look like is presented in figure 5.

5.2 The Collective Process of Data-Driven Decision Making

The collective process of DDDM was an emerging theme that relates to the need for a shared understanding of definitions in order to use DDDM in an organization. Still, this theme relates to multiple concepts discussed in the literature review such as the process of DDDM, technology and data, and organization and culture. Below, these connections will be further explored.

At its core, this theme relates to the need for a common business language that is built on clear and shared definitions. According to the interviewees, many decisions are the result of discussions. Some decisions are taken together, some decisions are influenced by stakeholders, and some decisions will affect stakeholders. Therefore, having clear and shared definitions creates a shared understanding of what the data and information really means.

According to Ikemoto and Marsh (2007), the transformation of data into information can be the result of either an individual process, or a collective process. In the collective process, data is collected and organized by multiple individuals in order to create information. Logically, all individuals who participated in the collective process of turning data into information should have a shared understanding of what it means. However, when the interviewees discussed definitions, they did not express a will to involve everyone in the company in the collection of data. Instead, they expressed a need of deciding on a set of definitions. As noted by participant 3, definitions should be agreed on by the ones who will use them the most. Then, the definitions should be shared with everyone else in order to create the common business language. Therefore, there is a difference to the collective process described by Ikemoto and Marsh (2007).

The need for clear and shared definitions also relates to Berntsson Svensson et al. (2019) findings that a reason to not use DDDM is that decision makers sometimes lack the necessary understanding of the data and how to use it. It is possible that their finding might relate to the need of clear definitions. As observed by participant 3, the lack of clear definitions could result in decreased trust in data, which reduces its usefulness. According to participant 3, this is due to people have different assumptions of what is meant by, for example, a newsletter registration. Thus, having clear and shared definitions is believed to reduce uncertainties regarding the information, and possibly increase the use of data in decision making.

Furthermore, the need for definitions also relates to technology and data. In the literature review, it was noted that knowledge about the data should be transferred to end users as it helps them understand how the data can be used (Janssen et al., 2017; Power, 2016). Likewise, participants 3 and 4 mentioned that reports should be transparent and include an explanation of what is measured and how. In addition, this also relates to that data should be trustworthy in order to be used in DDDM (Anderson, 2015). As mentioned by participant 3, without clear definitions it is possible that people will question if the information is correct. Thus, clear definitions could be seen as knowledge about the data which increases its trustworthiness.

A final connection can also be noted in regard to organization and culture where the need for a data-sharing culture was noticed (Anderson, 2015). On this topic, it was mentioned that low levels of trust between departments may result in data silos where people are unwilling to share information (Janssen et al., 2017). However, according to the participants, the reason for not sharing data is primarily that not all information can be readily understood by other departments or units. For example, participant 4 explains that while people would be willing to share, the obstacle lies in making the information relevant for others. Thus, a lack of clear definitions may be another obstacle to sharing data which was not mentioned in the reviewed literature.

However, Waller (2020) suggests that boundaries between business units and data scientists should be highly porous in order for business units to better understand data, while also allowing data scientists to better understand the business needs. This is similar to what participant 3 mentioned about including stakeholders in the creation of definitions. Therefore, close collaboration between units could facilitate the creation of clear and shared definitions.

Above, it has been explained how the emerging theme of a collective DDDM process relates to different parts of the reviewed literature. Although connections to multiple areas can be found, the collective process of DDDM has not been explored as a theme in itself in the reviewed literature. In contrast, this collective perspective was mentioned multiple times by participants 1, 2, 3, and 4 who believed it to be crucial for DDDM. Therefore, it was surprising that it had not been explored in more detail in the reviewed literature.

However, it should be noted that not all data-driven decisions need to be collective. In the interviews, several participants mentioned situations that can be understood as the individual processes described by Ikemoto and Marsh (2007). Here, it seemed that shared definitions were not important since the decision-maker made sense of their own data. For example, both participants 1 and 6 described that they have been analyzing data on their own. Furthermore, some DDDM processes may be collective but limited to a small group of people. This could for example be that there are clear and shared definitions within a business unit. On the other hand, participant 1 also explained that transparency towards stakeholders is important. Therefore, even when a decision was made through an individual DDDM process, it might be important to be able to motivate the decision with information that can be collectively understood.

5.3 Enabling Factors of Data-Driven Decision Making

5.3.1 Strategy

According to Davenport et al. (2001), DDDM is affected by a firm's strategy since it guides data collection. At Lynk & Co, participant 4 described that the end goal of DDDM is connected to the broader strategy by being a tool to help run the business better in terms of sustainability, creating experiences, and profitability. In addition, participants also perceived being data-driven as a strategic goal, albeit a broad one. This resulted in a number of different interpretations of how data should be used in decision making, which in turn resulted in variations of the data being collected between units.

Whereas the data collection described in the reviewed literature should be guided by a firm's vision and strategy (Davenport et al., 2001; McAfee & Brynjolfsson, 2012) so that leaders can select which metrics to focus on (Gupta & George, 2016), the DDDM at Lynk & Co appeared more decentralized. The space for interpretations described above allowed each unit to select metrics based on their own goals. For example, participant 6 described how they collect data based on the goal of driving conversion on the website. This is similar to what was recommended by Lin (2018) who suggested that it is better to select only the most relevant metrics instead of collecting more data than can be analyzed. In addition, participant 1 described that they need very specific data to guide their development. While the data collected by participant 1 is relevant to their unit's goal, it may not be relevant to another unit's goal. This highlights how the decentralized interpretations of what the strategy means to each unit allows for variations in the data being collected. As a result, each unit can collect data tailored to their needs, while being aligned with the overall strategy.

The findings above contributes to the literature by clarifying that a firm's strategic goals can be interpreted to guide the data collection of each unit. The firm's high level strategic goal does not appear to imply that only a specific set of metrics should be tracked across all units. Instead, each unit can collect the most relevant data for their specific goals that are connected to the overall strategy.

5.3.2 Skills and Experience

In the literature review, two types of skills related to DDDM were identified: technical skills, and managerial skills (Gupta & George, 2016). Technical skills were recognized as important both in turning data into information (Davenport et al., 2001), and also to understand what the produced information means (Provost & Fawcett, 2013). Therefore, the technical skill can be seen as two components that needs to be combined: the skill to turn data into information, and the skill to understand information.

According to the interviewees, the technical skill to turn data into information could reside in experts. Although most interviewees appeared to have the necessary skills, it was noted by participant 1 that not everyone needs to have it in order to engage in DDDM. Similarly, participant 6 mentioned that they would need a dedicated expert who could focus on data and bring information to the team. This is also similar to what was found in the literature where it was noted that the skill to turn data into information is most important to data scientists (Provost & Fawcett, 2013).

On the other hand, discussions about the technical skills to understand information gravitated towards the collective process of DDDM. Based on the literature, there should be a basic understanding of how to understand information across an organization (Davenport & Mittal, 2020; Provost & Fawcett, 2013). However, according to the interviewees, the understanding of information depends on how well KPIs have been defined and to what extent there is a common business language. Thus, even though everyone might have the basic technical skills, it cannot be leveraged unless there is a common business language with clear and shared definitions. Therefore, this links to the collective process of DDDM discussed in 5.2.

In terms of managerial skills, Janssen et al. (2017) explains that when decision-makers start using DDDM, they are likely to experience uncertainty regarding how data should be used in their decisions, and if the decisions are correct. However, this uncertainty was not noted among the participants. Instead, several participants expressed that having data makes them feel more comfortable with their decisions. For example, participant 1 expressed that data makes decisions more transparent towards stakeholders. Moreover, participant 4 mentioned that DDDM allows for a better follow-up on decisions. Therefore, unlike what was found in the literature, participants expressed that they would feel more confident with their decisions if they were based on data.

While it is possible that the participants selected for this study may have more favorable opinion towards data than others, their arguments for why data can decrease uncertainty for decision-makers are still relevant to a broader audience. Whereas Janssen et al. (2017) describes that decision-makers may feel insecure when increasing their use of data in decision making, the participants of this study makes a compelling argument on why data can also increase the confidence with decisions. Thus, it is possible that increased DDDM may cause decision-makers to experience decreased, and increased, confidence with their decisions at the same time. The decreased confidence would according to Janssen et al. (2017) be due to a new, unfamiliar, way of decision making. On the other hand, the transparency and ability to follow up on decisions could increase confidence.

5.3.3 Technology and Data

According to the reviewed literature, DDDM cannot be performed without data that is trustworthy, timely, and accurate (Anderson, 2015). Nor can it be performed if the decision maker does not have access to data (Anderson, 2015; Waller, 2020). Based on the interviews, the trustworthiness of data appeared, to a large degree depend on the collective process of DDDM. This is discussed in greater detail in 5.2.

Data access on the other hand was recognized as important both in the literature (Anderson, 2015; Waller, 2020) and by participants 5 and 7. Whereas data access as described by Waller (2020) appears to mainly focus on the technical issues of providing decision makers with internal data, participants 5 and 7 highlighted that data access is also related to contracts and relationships with other companies. On this topic, participants 5 and 7 mentioned two instances in which data access is dependent on external companies. First, some systems that generate data is owned by other companies. Then, data access is limited to what the other company is willing, or capable of sharing. Even though participant 5 believed that their suppliers have access to data but were not willing to share all of it, another scenario could possibly be that other companies are missing the needed capabilities to collect, organize and share data. Second, data access can also be limited when data collection is not included as a requirement when a product is being co-developed with another company. Hence, participants 5 and 7, agrees that data access should be discussed with all external partners from the start.

However, it should be noted that figure 2 by Jia et al. (2015) includes supply chain partners as a stakeholder affecting the DDDM process. Still, apart from stating it, they offer little explanation of how the DDDM process is affected by suppliers and what effects it has. Therefore, the participants offer additional insights that complements this figure.

Based on the interviews, the topic of technology and data can thus be expanded to also include external companies. While access to high quality internal data is still important, this addition underlines that it is also related to contracts and relationships with other companies. Since, as stated by participant 5, dependencies on other companies are common, this is a perspective that should be considered.

5.3.4 Organization and Culture

According to the reviewed literature, organizational culture is seen as one of the key enablers of DDDM (Davenport & Mittal, 2020; McAfee & Brynjolfsson, 2012) since it encourages decision makers to notice, trust and use data (Anderson, 2015). During the interviews, it became clear that all participants had positive attitudes towards data in their decision making. However, it was also noted that this emphasis on data was part of an ongoing cultural shift.

As described by participants 3, 5, and 6, the movement from a start-up to scale-up has resulted in a shift from making decisions based on intuition, to making decisions based on data. From the interviews, it is clear that this shift is the result of more data now being available, the new strategic goal of DDDM, and that people in the company has been requesting more data. Overall, participants described that data was being requested both to increase transparency of decisions, and also to be the basis of decisions. That data is being requested shows that people are expecting data to be used in decision making. This indicates a data-driven culture according to Davenport and Mittal (2020) and Waller (2020). Additional similarities to the literature regarding data-driven cultures could also be found based on the interviews. Thus, the description of data-driven cultures in the reviewed literature was, at large, found to be the case in this study as well.

However, as mentioned by participant 6, this cultural shift was ongoing which meant that not all departments had made the same advancements. While some product development was taking place in settings where the data-driven culture was shared by everyone included, participant 6 mentioned that some of their stakeholders did not have the same view on data in decision making. As a result, stakeholder 6's team and the stakeholders arrived at different priorities, which caused disagreements. This result is similar to what was described in the reviewed literature. According to Davenport et al. (2001) a data-driven culture needs to include the entire organization. If that is not the case, Carillo (2017) explains that friction between employees may occur.

Due to the focus on product development, interviews with the stakeholders were not conducted. Therefore, it is not known what caused them to have less positive attitude towards data. This meant that the discussions by Brynjolfsson and McElheran (2016) and McAfee and Brynjolfsson (2012) regarding how the tenure of employees and HiPPO were not explored as potential reasons. However, participant 3 believed that most people would want to use data, but that the lack of clear and shared definitions caused uncertainty which negatively affected the attitudes towards using data. This was interesting since a data-driven culture, according to Anderson (2015), is characterized by a shared belief that data can be trusted. Here, it could be possible to assume that the degree to which people trust data depends on the data itself, and if it is defined clearly enough. Therefore, the development of a data-driven culture could also depend on the collective process of DDDM.

Apart from data-driven cultures, the reviewed literature also mentioned data-sharing cultures. According to Janssen et al. (2017), it is important to enable collaboration and knowledge sharing in an organization by building trust between departments. If this is not successful, it may result in data silos which hinders DDDM (Anderson, 2015). According to the interviewees, information was being shared both through reports and ad hoc by asking people for the needed information. However, participants 1, 4, 6, and 7 also mentioned that they did not feel a need to share or receive all information. This was not understood as a sign of low trust between departments, but rather that they felt that all data was not relevant to share. In addition, participant 4 also mentioned that sharing all data would require time to make it relevant for others. Likewise, participant 7 mentioned that sharing conclusions might be more relevant than sharing the data itself. Therefore, a data-sharing culture may not require that all data is being shared through reports. Instead, the ad-hoc sharing by asking others for information may be more suitable for certain information and data.

5.3.5 Basic Resources

While the reviewed literature was rather sparse in terms of basic resources other than mentioning that time and money needs to be invested (Gupta & George, 2016), the participants expanded this category by explaining that time is needed to turn data into information, and to read the produced reports. In terms of turning data into information, this also relates to the collective process of DDDM where time is needed to discuss and agree on definitions. To read reports, on the other hand, participants mentioned that they are sometimes too busy to give information the necessary attention. This indicates that while time and money are needed as an initial investment in DDDM (Brynjolfsson & McElheran, 2016; Gupta & George, 2016), it also requires ongoing investments by ensuring that everyone has the time needed to internalize information into knowledge.

Lastly, it was also mentioned by participant 2 that including more people could result in less robust solutions when building tools and KPIs for DDDM. This was hypothesized due to the increasing challenges of coordination when more people are included. According to Brynjolfsson and McElheran (2016), investments in IT is positively correlated with DDDM adoption. However, based on the interview with participant 2, it could be possible to assume that there is a decreasing marginal effect of involving more people when building solutions for DDDM.

5.4 Data-Driven Decision Making Based on Assumptions

The issue of using historical data for new products has been mentioned both by the reviewed literature and the interviewees. Martin and Golsby-Smith (2017) argues that DDDM based on historical data will impede innovation which makes it unsuitable when creating something completely new. Similarly, this limitation of historical data was also noted by participants. As explained by participants 3 and 4, DDDM in this context rather revolves around gathering data to test assumptions than looking at historical data. This is similar to what Martin and Golsby-Smith (2017) described as data creation through prototyping.

According to Davenport (2013), this testing of assumptions, or hypotheses, is what separates the data-driven from acting solely on intuition. This was also noted by participant 4 who mentioned that deciding based on assumptions is fine, but that the goal should be to follow up on the assumptions to see if they were true. Therefore, both the reviewed literature and the participants agree that new developments may require new data to be generated, through for example prototyping, in order to test assumptions and hypotheses.

Therefore, it appears that DDDM in this context is somewhat different from the DDDM process described in 2.2 which starts with data. Instead, both the reviewed literature and the participants mentions assumptions and hypotheses as starting points. Still, once an assumption has been made and data has been generated and collected, the DDDM process described in 2.2 becomes applicable again. Therefore, there appears to be some connection to the process of DDDM, though the exact connection remains ambiguous. However, it is also possible that these are separate, but complementary, processes. This would be similar to how Lichtenthaler (2018) considers the relationship between human and AI to be complementary in creative work. Still, this is an area that could benefit from further research.

Regardless, the interviews together with the reviewed literature suggests that there is a space for data-driven decision making even where there is no historical data. Both the reviewed literature and the interviewees agree that assumptions can be made, as long as they are tested to inform future decisions. Therefore, the findings of this thesis do not suggest that DDDM will hinder creativity and innovation.

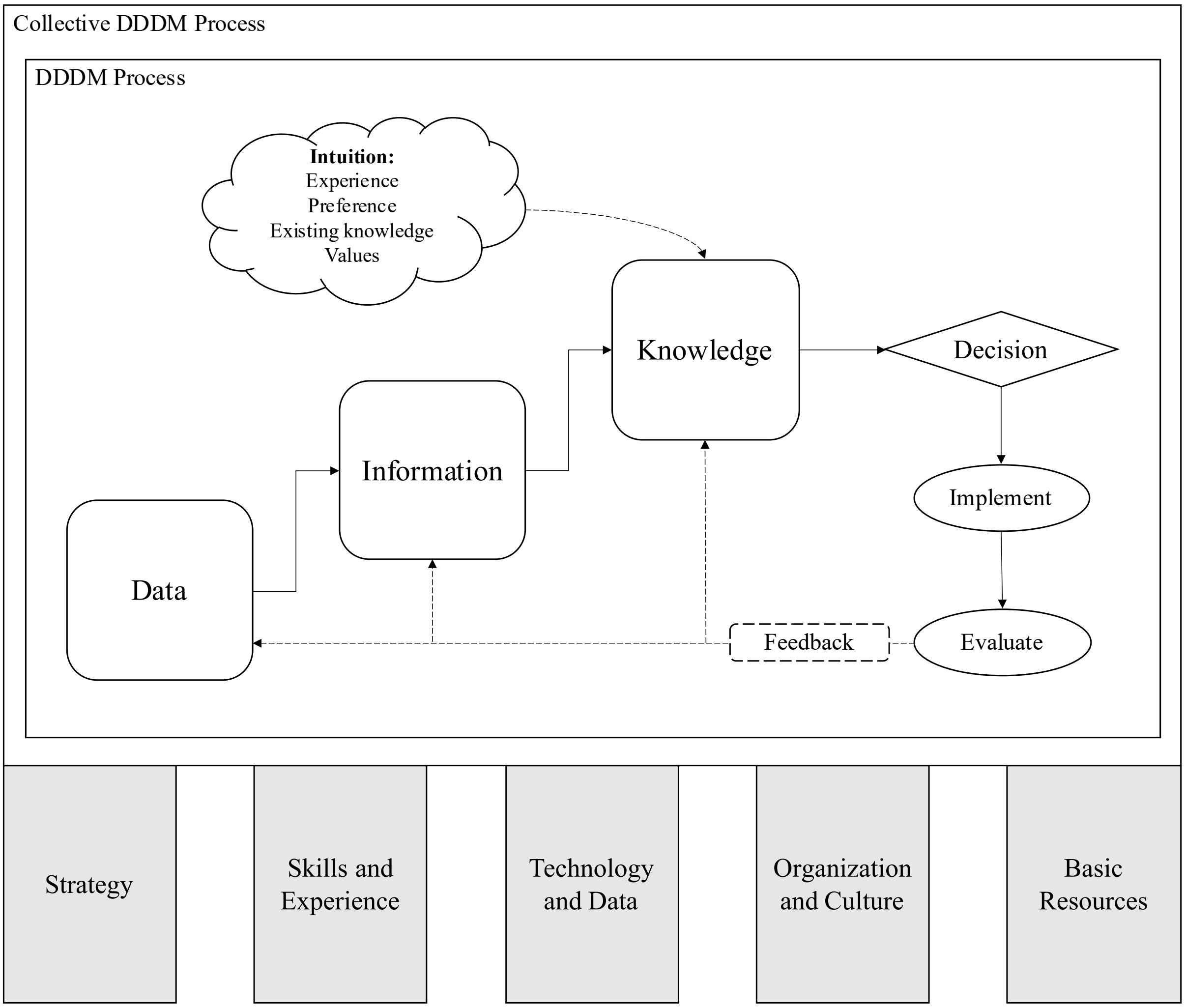
5.5 Key Takeaways

In this chapter, the results from the interviews have been compared with the theoretical framework created in the literature review. As a result, several similarities and differences have been identified. Although not all parts of the literature review have been mentioned in this chapter, they are still considered relevant. Instead, this chapter has focused on highlighting how the interviewees' perception of DDDM compares to what has been mentioned in the reviewed literature.

Based on these contributions to the reviewed literature, a new model is proposed in figure 5. This model is based on figure 3 but adds an intuition component and the collective process of DDDM. Although contributions have also been made to the enabling factors, they are not visible in the model itself. This model shows how both data and intuition results in knowledge that influences a decision. However, it is important to keep in mind the definition of DDDM as "*the practice of basing decision on the analysis of data rather than purely on intuition*" (Provost & Fawcett, 2013, p. 53). Therefore, the dashed line from intuition indicates that data should have a greater influence on the decision. Furthermore, the collective process of DDDM is closely connected to multiple enabling factors. It could also be understood that although an individual process of DDDM can take place, decisions are likely to affect stakeholders. Therefore, the collective process of DDDM has been added around the DDDM process in connection to the enabling factors.

Figure 5

(Collective) DDDM Process and Enabling Factors



Note. Figure produced by the author

6. Conclusion

In this chapter, the main findings will be presented in order to answer the research question. This will be followed by recommendations for organizations that wants to increase their use of data-driven decision making, as well as recommendations for future research. In this chapter, limitations of the study will also be discussed.

6.1 Answering the Research Question

6.1.1 Purpose and Research Question

This thesis started by recognizing that companies are struggling to leverage data in their decision making. Despite investments in data initiatives, few companies are able to succeed with their attempts to become data-driven (Bean, 2021; Bean & Davenport, 2019). While plenty of attention has been given to the technical aspects of data-driven decision making, it was noted that there is a need for additional knowledge regarding the organizational perspectives of it (Berntsson Svensson et al., 2019; Gupta & George, 2016; Mikalef et al., 2018). Therefore, the aim of this thesis was to explore how data-driven decision making can be used within product development. Based on this, the following research question was formulated:

To what extent can data-driven decision making be used in product development?

6.1.2 Main Findings

The research question has been answered through both a literature review and a case study with semi-structured interviews. With the theoretical framework as a starting point, the case study has resulted in several contributions to existing literature. As previously noted, there has been a research gap regarding the people perspective of data-driven decision making. This gap has been addressed in this study by focusing on people's perception of working with data in product development. The main theoretical contributions are described below.

First, it was discovered that since there is room for intuition according to the definition of data-driven decision making (Provost & Fawcett, 2013), the same data may result in different decisions depending on the decision-maker. Experience, preference, existing knowledge, and values appears to affect how information is internalized into knowledge. Consequently, this should be added to the process of data-driven decision making as suggested in figure 5. In addition, it should also be noted that this intuition component is individual. Therefore, when multiple decision-makers have different values, they are likely to prefer different options. Consequently, this contributes to the theoretical understanding of data-driven decision making by clarifying that individual intuition is likely to affect the process of data-driven decision making.

Second, a new theme emerged from the interviews and should be included in the understanding of data-driven decision making. This theme has been referred to as the collective process of data-driven decision making. It is grounded in the participants' belief that decisions in product development are not taken in isolation. Either a decision is taken by a group, or it is affected by, or affecting, stakeholders. Therefore, a common business language built on clear and shared definitions is needed to ensure efficient communication, increase trust in data, and to enable information sharing. Based on the interviews, this theme appeared to be closely connected to multiple enabling factors. Therefore, the collective process of data-driven decision making has been added to figure 5 in connection to all enabling factors in a way that it encompasses the process of data-driven decision making. Thus, this contribution to theory provides a previously unexplored aspect of data-driven decision making that is likely to affect the extent to which it can be used in product development.

Third, several contributions were made to the enabling factors of data-driven decision making. Here, it was noted that (1) broad strategic goals can be interpreted into unit specific goals which helps guiding data collection. (2) data-driven decision making might have a positive effect on managerial skills since being able to follow up on transparent decisions appears to increase confidence among interviewees. (3) Technology and data can be expanded to also include relationships and contracts with external partners and suppliers since they may restrict data access. (4) A data-driven culture does not necessarily develop at the same rate across a company, which can cause friction between departments. (5) Time as a resource is needed to turn data into information, but equally important, to internalize information.

Finally, it was also noted that data-driven decision making can, to some extent, be used when developing completely new products where there is no historical data. However, it appears that this decision-making process starts with assumptions and hypotheses rather than data. Still, these assumptions are tested by collecting data to learn more about the decision that was made. Thus, it guides future decisions towards the process of data-driven decision making where data is the basis for the next decision.

Taken together with the initial theoretical framework, these findings suggests that data-driven decision making can be used in product development to a certain extent. However, to what extent it can be used appears to depend on: (1) how well the process of data-driven decision making gets implemented, (2) if a collective effort is made to ensure a common business language with clear and shared definitions, and (3), to what extent the enabling factors are present.

6.2 Recommendations

In the first chapter, it was noted that despite companies' investments to become data-driven, few succeeds with their attempts (Bean, 2021; Bean & Davenport, 2019). Furthermore, it was noted that companies need knowledge regarding how their data and technology can be leveraged in their decision-making process. Based on this thesis, three recommendations will be provided to organizations that wants to increase their use of data-driven decision making in product development.

First, it is important to consider the collective process of DDDM. To establish a common business language with clear and shared definitions is crucial as it affects trust in data, information sharing, and facilitates discussions. However, to arrive at company-wide definitions in all areas may be a lengthy project. Therefore, it could be suggested to start small with the most important definitions. For example, this can be done in cross functional meetings where data is being presented and clarified. According to McAfee and Brynjolfsson (2012), this helps establish a data-driven culture where participants have access to the same information. Also, these definitions should be included in all reports in order to increase trust in data.

Second, it is important to manage the data-driven culture across the organization. When technical and cultural progress may happen at a faster rate in certain departments, disagreements may occur. This is also related to the common business language where uncertainty in data may vary between departments. Therefore, initiatives in data-driven decision making in product development should be expanded to also include the relevant stakeholders. For example, this could be done by including stakeholders in the formulation of goals that can be tracked using KPIs. This should, according to Brynjolfsson and McElheran (2016) motivate action towards data-driven decision making.

Third, it is important to consider relationships and contracts with other companies as this might affect data access. While this recommendation is less applicable for in-house developments, it is important to consider that data-access is not only a technical issue, but also connected to the data that can be received from suppliers and partners.

6.3 Limitations

While the chosen research strategy and design resulted new knowledge, it should be noted that it came with its limitations. First, this case study focused primarily on data-driven decision making at the Business Technology department at Lynk & Co. However, in order to get a more balanced view on the topic, other departments from the case company could have been involved. For example, several interviewees discussed decision making in relation to their stakeholders. Thus, a limitation of this study was that only one group of participants was interviewed. Therefore, the study could have been improved if more groups would have been added.

A second limitation of this study is related to the chosen research design. Whereas a single case study has its advantages, its limitation is the relatively low transferability and weak external validity (Bell et al., 2019). Instead, a cross-sectional design of at least two case companies could have resulted in a wider understanding of data-driven decision making not limited to the context of product development of software intensive products. This would have allowed for a better understanding of the variation between different companies, and perhaps industries, in regard to data-driven decision making (Bell et al., 2019).

A third limitation of this study, related to the chosen research strategy, is that only one method was used. In order to increase the confidence of the study through triangulation, mixed methods could have been used. For example, this could have been done by adding a quantitative component after the qualitative one (Bell et al., 2019). For this study, this would have meant that a questionnaire could have been produced after the interviews to investigate the transferability of the findings.

6.4 Further Research

Data-driven decision making is a research topic still under development and with increased interest among organizations, it is a topic that could benefit from future research. During this thesis, a number of different concepts related to data-driven decision making has been explored, but there is still plenty left to discover. Therefore, the following research areas could benefit from increased knowledge.

First, it was noted that a cultural shift towards data-driven decision making may happen at different rates in an organization. Similarly, the maturity relating to data-driven decision making could vary between different departments. Therefore, it would be interesting to further develop the model proposed in this study by turning it into a maturity model. Since it is likely that the enabling factors develop at different rates, this could also serve as a tool for companies to guide their efforts in becoming data-driven. Furthermore, such a tool could also be used to track progress across multiple departments.

Second, whereas this study focuses on data-driven decision making where a human decision maker is present, it should be noted that technical advancements increasingly allow for automated data-driven decision making. This means that certain decisions are made by a computer instead of a human (Provost & Fawcett, 2013). Therefore, it would be interesting to understand how automated data-driven decision making relates to the proposed model in this study which emphasized the people perspective.

Third, this study's research setting was to understand data-driven decision making within software product development in a business technology department. Since this process may be different compared to data-driven decision making in other settings, this study could be replicated in other industries as well. Alternatively, a cross-sectional research design could allow for comparisons of two companies using data-driven decision making in different industries. On the other hand, a cross-sectional design could also be used to compare two similar companies that are at different stages of implementing data-driven decision making. This comparison could allow for a better understanding of the necessary enabling factors.

References

Anderson, C. (2015). *Creating a data-driven organization: Practical advice from the trenches*. O'Reilly Media, Inc.

Bean, R. (2021). *Why Is It So Hard to Become a Data-Driven Company*. Retrieved 2021-12-23 from https://hbr.org/2021/02/why-is-it-so-hard-to-become-a-data-driven-company

Bean, R., & Davenport, T. H. (2019). *Companies are failing in their efforts to become data-driven*. Harvard Business Review. Retrieved 2022-03-08 from https://hbr.org/2019/02/companies-are-failing-in-their-efforts-to-become-data-driven

Bell, E., Bryman, A., & Harley, B. (2019). *Business research methods* (Fifth ed.). Oxford : Oxford University Press.

Berntsson Svensson, R., Feldt, R., & Torkar, R. (2019). The unfulfilled potential of data-driven decision making in agile software development. 20th International Conference, XP 2019, Montréal, QC, Canada.

Bhansali, N. (2013). *Data governance: Creating value from information assets*. CRC Press.

Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, *70*, 328-337.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decisionmaking affect firm performance? *Available at SSRN 1819486*.

Brynjolfsson, E., & McElheran, K. (2016). The Rapid Adoption of Data-Driven Decision-Making. *American Economic Review*, *106*(5), 133-139. https://doi.org/10.1257/aer.p20161016

Carillo, K. D. A. (2017). Let's stop trying to be "sexy" – preparing managers for the (big) data-driven business era. *Business process management journal*, *23*(3), 598-622. https://doi.org/10.1108/BPMJ-09-2016-0188

Chen, C. L. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314-347. https://doi.org/10.1016/j.ins.2014.01.015

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4), 1165-1188. https://doi.org/10.2307/41703503

Davenport, T. H. (2013). *Big Data and the Role of Intuition*. Harvard Business Review. Retrieved 2022-03-15 from https://hbr.org/2013/12/big-data-and-the-role-of-intuition

Davenport, T. H., Harris, J. G., De Long, D. W., & Jacobson, A. L. (2001). Data to Knowledge to Results: Building an Analytic Capability. *California management review*, *43*(2), 117-138. https://doi.org/10.2307/41166078

Davenport, T. H., & Mittal, N. (2020). *How CEOs Can Lead a Data-Driven Culture*. Harvard Business Review. Retrieved 2022-02-27 from https://hbr.org/2020/03/how-ceos-can-lead-a-data-driven-culture

Desjardins, J. (2019). *How much data is generated each day?* . World Economic Forum. Retrieved 2022-03-09 from https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/

Elgendy, N., Elragal, A., & Päivärinta, T. (2021). DECAS: a modern data-driven decision theory for big data and analytics. *Journal of decision systems*, 1-37. https://doi.org/10.1080/12460125.2021.1894674

European Training Foundation. (2018, December 16 2018). *'Without data, you're just another person with an opinion', W. Edwards Deming*. European Training Foundation. Retrieved 2022-03-09 from https://www.etf.europa.eu/en/news-and-events/news/without-data-youre-just-another-person-opinion-w-edwards-deming

Flyvbjerg, B. (2010). Five Misunderstandings About Case-Study Research. *Qualitative inquiry*, *12*(2), 219-245. https://doi.org/10.1177/1077800405284363

Goffin, K., Åhlström, P., Bianchi, M., & Richtnér, A. (2019). Perspective: State-of-the-Art: The Quality of Case Study Research in Innovation Management. *The Journal of product innovation management*, *36*(5), 586-615. https://doi.org/10.1111/jpim.12492

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064. https://doi.org/https://doi.org/10.1016/j.im.2016.07.004

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the "Data-Driven" Mantra: Different Conceptions of Data-Driven Decision Making. *Teachers College Record*, *109*(13), 105-131. https://doi.org/10.1177/016146810710901310

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, *70*, 338-345. https://doi.org/https://doi.org/10.1016/j.jbusres.2016.08.007

Jia, A., Hall, D., & Song, J. (2015). *The Conceptualization of Data-driven Decision Making Capability The Conceptualization of Data-driven Decision Making Capability* Twenty-first Americas Conference on Information Systems, Puerto Rico.

Lichtenthaler, U. (2018). Substitute or Synthesis: The Interplay between Human and Artificial Intelligence. *Research technology management*, *61*(5), 12-14. https://doi.org/10.1080/08956308.2018.1495962

Lin, C. (2018). Data driven product management. *IEEE engineering management review*, *46*(1), 16-18. https://doi.org/10.1109/EMR.2018.2810099

Lynk & Co. (2021). *Cheers to five years: Lynk & Co celebrates anniversary with unprecedented Month-to-Month membership milestone*. Retrieved 2022-03-08 from https://press.lynkco.com/en-WW/203553-cheers-to-five-years-lynk-co-celebrates-anniversary-with-unprecedented-month-to-month-membership-milestone

Mandinach, E. B. (2012). A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice. *Educational psychologist*, *47*(2), 71-85. https://doi.org/10.1080/00461520.2012.667064

Mandinach, E. B., Honey, M., & Light, D. (2006). A theoretical framework for data-driven decision making. Annual meeting of the American Educational Research Association, San Francisco, CA.

Martin, R. L., & Golsby-Smith, T. (2017). Management Is Much More Than a Science. *Harvard business review*, 128-135.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, *90*(10), 60-128.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, *16*(3), 547-578. https://doi.org/10.1007/s10257-017-0362-y

Nair, S. (2020). *Is Your Business Masquerading as Data-Driven?* Harvard business review. Retrieved 2021-12-23 from https://hbr.org/2020/05/is-your-business-masquerading-as-data-driven

Power, D. J. (2016). Data science: supporting decision-making. *Journal of decision systems*, *25*(4), 345-356. https://doi.org/10.1080/12460125.2016.1171610

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big data*, *1*(1), 51-59. https://doi.org/10.1089/big.2013.1508

Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You May Not Need Big Data After All. *Harvard business review*, *December, 2013*.

Siggelkow, N. (2007). Persuasion with Case Studies. *Academy of Management journal*, *50*(1), 20-24. https://doi.org/10.5465/AMJ.2007.24160882

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Treder, M. (2019). *Becoming a Data-Driven Organisation: Unlock the Value of Data*. Berlin, Heidelberg: Springer Berlin / Heidelberg.

Waller, D. (2020). *10 Steps to Creating a Data-Driven Culture*. Harvard Business Review. Retrieved 2022-02-23 from https://hbr.org/2020/02/10-steps-to-creating-a-data-driven-culture

Appendices

Appendix 1: Key Takeaways from the Literature Review

Elements	Key Takeaways
The Process of Data-Driven Decision Making	

Data	• Data needs to be collected (Mandinach et al., 2006)
	  ○ Data exists in a raw form that has little meaning in itself (Mandinach et al., 2006).
	  ○ Data needs to be trustworthy, unbiased, timely, and accessible (Anderson, 2015).
	  ○ Data might need to be created through experiments (Martin & Golsby-Smith, 2017).
	• Data needs to be organized (Mandinach et al., 2006)
	  ○ The collected data is organized in a way that enables further interpretation (Mandinach, 2012).

Information	• Data put in context gives information (Mandinach et al., 2006) • Can be done through reporting, alerts, and analysis (Anderson, 2015).
Knowledge	• The decision-maker internalizes the information through synthesis and turns it into knowledge that gets prioritized (Mandinach, 2012; Mandinach et al., 2006).
Decision	• A decision is made based on the knowledge created (Mandinach et al., 2006). • The decision results in an impact that is evaluated and feeds back into the process. This makes it an iterative process (Jia et al., 2015; Mandinach et al., 2006)
Enabling Factors	
Strategy	• A clear vision, goals, and a definition of success guides the use of data in an organization (McAfee & Brynjolfsson, 2012). • Strategy guides data collection (Davenport et al., 2001). ○ Picking metrics that are relevant to the decision maker and tells a cohesive story facilitates decision making (Lin, 2018). • Strategy creates organizational support (Davenport et al., 2001). • Strategy motivates action in response to data (Brynjolfsson & McElheran, 2016).
Skills and Experience	• Skill and experience are needed to turn data into knowledge and decisions (Davenport et al., 2001). • Technical skills are needed to transform data into information (Davenport et al., 2001). ○ This skill is especially important for data scientists (Provost & Fawcett, 2013).

	o Basic technical skills are needed throughout the organization in order to understand the analyses (Provost & Fawcett, 2013). • Managerial skills are needed to use the information produced (Gupta & George, 2016). o Over time, decision makers get more experience of using DDDM which makes them more confident in relying on data (Janssen et al., 2017).
Technology and Data	• Data needs to be trustworthy, accurate and timely (Anderson, 2015). • Most common barrier to using data in decision making is that there is too little, or too much data available (Berntsson Svensson et al., 2019). o With too little data available, it is crucial that data access is provided to users (Anderson, 2015). o With too much data available, a better strategy for data collection might be needed (Lin, 2018). o Visualizations and storytelling can make data more understandable (Lin, 2018).
Organization and Culture	• Organizational culture encourages decision makers to notice, trust, and use data (Anderson, 2015). • Culture is often the ultimate obstacle to DDDM (Anderson, 2015; Bean & Davenport, 2019; McAfee & Brynjolfsson, 2012). • A culture that supports DDDM: o Has a shared belief that data can be trusted and should be used in decision making (Anderson, 2015). o Values, and expects decisions to be based on data (Davenport et al., 2001; Davenport & Mittal, 2020; Waller, 2020). o Trusts data more than intuition (McAfee & Brynjolfsson, 2012). o Has high degrees of collaboration and data-sharing between departments (Anderson, 2015; Janssen et al., 2017).
Basic Resources	• Basic resources, such as time and money, needs to be invested in a DDDM initiative (Gupta & George, 2016).

Appendix 2: Interview Guide

Area	Questions
Background	
General	• Could you start by telling me for how long you have worked at Lynk & Co? • Can you tell me more about your role at Lynk & Co? • What type of decisions do you make in your role?
Attitudes towards Data	• How do you normally make decisions in your daily work? • Do you think data can help you make better decision?
DDDM Process	
Data	• To what extent would you say that data is being collected? • Do you think that the data that is collected is useful to you? ○ Trustworthy, timely, accurate?
Information	• In what way is data presented to you? ○ Reports, excel files, raw data? • Do you make any further analysis of the data?
Knowledge	• When making a decision, do you feel that you understand the situation better when you get data/information? • Do you get the data/information you need to increase your knowledge?
Decision	• Do you base your decisions mostly on data or on intuition? • Think of hypothetical decision you need to make. In a perfect world, how would you use data or intuition to decide? • How is that different from your situation today?
Enabling Factors	
Strategy	• Do you think the Lynk & Co strategy includes making decisions based on data? • How would you describe what the goal of using data? • Are you aware of any targets regarding the use of data? • Do you feel that the strategy guides how you should use data? ○ What data/information to use?

Skills and Experience	• Do you think that most people working here knows how to analyze data?
	  ○ Is that also how you feel?
	• Do you think that most people working here has the basic skills needed to understand analyses that others have produced?
	  ○ Is that also how you feel?

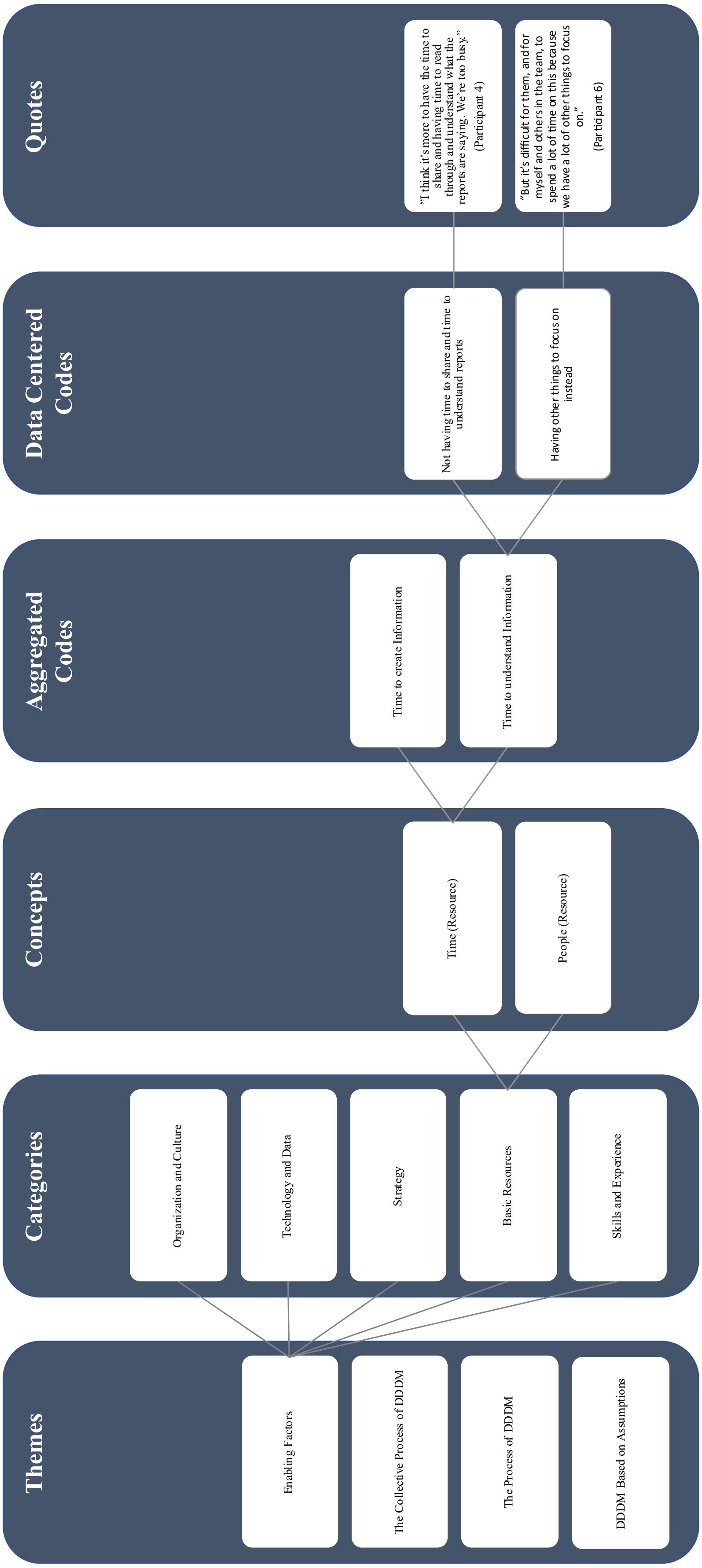
Technology and Data	• Do you feel that you have data available to you when you need it?
	  ○ Data or information?
	  ○ Too little data?
	  ○ Too much data?
	• In an ideal world, what data or information would you want?
	• How is that different from today?

Organization and Culture	• How would you describe the attitudes towards data in decision making at Lynk & Co?
	  ○ Do you feel that people trust and use data/information?
	  ○ Do you feel that decisions based on data are appreciated?
	• Do you think that data/information or intuition matters most when making a decision?
	  ○ Does it matter if someone influential has an intuition that is opposite of what the data says?
	• How would you describe the collaboration within your team when it comes to sharing data, information, and knowledge?
	• How would you describe the collaboration between departments when it comes to sharing data, information, and knowledge?
	  ○ Do you share data from your department with other collogues?

Basic Resources	• How would you describe the investments that are made towards using data in decision making?
Concluding Questions	
Concluding Questions	• If you would advise Lynk & Co on using more data in decision making, what would your recommendations be? • Is there anything else you would like to add about this topic?

Appendix 3: Thematic Analysis

Coding Example



Thematic Framework

Themes	Categories	Files	References
Enabling Factors		7	172
	Organization and Culture	7	65
	Technology and Data	7	58
	Strategy	6	26
	Basic Resources	5	24
	Skills and Experience	6	9
The Collective Process of DDDM		4	52
	Definitions are needed	4	37
	Creating Definitions	2	15
The Process of DDDM		7	41
	Knowledge and Decision	6	21
	Iterative Process	4	10
	Data	5	6
	Information	3	4
DDDM Based on Assumptions		2	4