Antti Kallio

A MULTI-CASE STUDY OF THE ANALYTICAL CAPABILITIES OF FINNISH E-COMMERCE BUSINESSES – A RESOURCE-BASED VIEW

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2021

TIIVISTELMÄ

Kallio, Antti
Monitapaustutkimus suomalaisten verkkokauppojen analytiikkakyvykkyyksistä – resurssiperusteinen näkökulma
Jyväskylä: Jyväskylän yliopisto, 2021, 63 s.
Tietojärjestelmätiede, pro gradu- tutkielma
Ohjaaja: Frank, Lauri

Datan määrän eksponentiaalisen kasvun myötä data-analytiikasta on tullut yhä keskeisempi tekijä organisaation päätöksenteossa. Data-analyysien pohjalta yrityksillä on käytössään enemmän tietoa kuin koskaan ennen ja päätökset pohjautuvat yhä enemmän analytiikan tarjoamaan informaatioon. Data-analyysien käytön mahdolliset liiketoiminnalliset hyödyt ovat selvät, mutta data-analytiikka vaatii kuitenkin yritykseltä tiettyjä kyvykkyyksiä, jotka voidaan jaotella aineellisiin, inhimillisiin sekä aineettomiin tekijöihin. Tämän opinnäytetyön tarkoituksena on tuottaa laaja ja selkeä kuvaus analytiikan hyödyntämiseen tarvittavista kyvykkyyksistä. Kirjallisuuskatsauksen tulosten perusteella onnistuttiin tunnistamaan kaikki keskeiset tekijät, jotka liittyvät yrityksen analytiikan hyödyntämiseen. Kirjallisuuskatsauksessa todettiin yhteenvetona, että yritys tarvitsee kyvykkyyksiä jokaisella osa-alueella voidakseen hyödyntää analytiikkaa omassa toiminnassaan. Kirjallisuuskatsauksen lisäksi tehtiin kymmenen puolistrukturoitua haastattelua alan ammattilaisten kanssa kuvaamaan suomalaisten verkkokauppayritysten nykytilaa. Empiirisessä osassa tutkielmaa pyrittiin kirjallisuuskatsauksen avulla selvittämään, mitkä näistä tunnistetuista kyvykkyyksistä korostuvat ja millä osa-alueilla suomalaisissa verkkokauppayrityksissä havaitaan puutteita. Tulokset osoittavat, että kirjallisuuskatsauksen kymmenen tunnistettua resurssia ovat hyvin linjassa haastateltavien vastausten kanssa, lukuun ottamatta tiedonhallintaan liittyviä asioita, joita ei tullut haastatteluissa esiin. Tärkeimmät haastateltavien tunnistamat ongelmat olivat teknisten taitojen alhainen taso sekä vaikeus löytää työntekijöitä, joilla on hyvä ymmärrys liiketoiminnasta ja hyvä ymmärrys analytiikasta sekä sen kehittämisestä. Lisäksi haastatteluissa kävi ilmi, että dataan perustuva päätöksenteko on yrityksissä usein vajavaista. Lisäksi korostettiin, että yrityksen sisällä useampien tulisi ymmärtää paremmin dataa ja analytiikkaa. Materiaalisia tekijöitä, kuten rahaa, aikaa, dataa ja verkkokauppa-alustoja, pidettiin enimmäkseen neutraaleina tekijöinä, eikä niitä pidetty ongelmana analytiikan kehittämisessä.

Asiasanat: analytiikka, verkkokauppa, kyvykkyys

ABSTRACT

Kallio, Antti
A multi-case study of the analytical capabilities of Finnish e-commerce businesses – a resource-based view
Jyväskylä: University of Jyväskylä, 2021, 63 pp.
Information Systems, Master's Thesis
Supervisor(s): Frank, Lauri

With the exponential growth in the amount of data, data analytics has become an increasingly central factor. Based on data analysis, organizations have more information at their disposal than ever before, and decisions are based more on the information provided by analytics rather than intuition. The potential business benefits of using data analytics are clear, but data analytics still requires a company to have certain capabilities that can be divided into tangible, human and intangible resources. The purpose of this thesis is to produce a wide and clear classification of the capabilities required to utilize data analytics. The classification has been carried out on the basis of previous literature. Literature review's results managed to identify broadly all the key factors involved in leveraging the company's analytics. The literature review summarized that a company needs capabilities in every subarea in order to be able to leverage analytics in its own operations. In addition to the literature review, ten semi-structured interviews with industry professionals were conducted to describe the current situation of Finnish e-commerce businesses. In the empirical part, the aim was to find out, with the help of a literature review, which of these identified capabilities are emphasized and in which areas deficiencies in Finnish e-commerce companies are identified. The results show that the ten identified resources of the literature review are well aligned with the interviewees' responses, except for governance, which did not come up in the interviews. The main problems identified by interviewees was the low level of technical skills and the difficulty in finding employees with a good understanding of the business and a good understanding of analytics and how to develop them. In addition, the interviews revealed a lack of data-driven decision making. The importance of everyone having some understanding of data was also highlighted. Material factors such as money, time, data and e-commerce platforms were mostly seen as neutral factors and were not seen as a problem in the development of analytics.

Keywords: analytics, e-commerce, capability

FIGURES

TABLES

TABLE OF CONTENTS

1 INTRODUCTION

E-commerce companies operate in a rapidly changing and growing business environment. According to recent surveys, including last year's Posti's (2020) survey, almost sixty percent of Finns buys something online every month and almost a third every week. This recently made study also shows the effect of corona pandemic to consumers' behaviour. Posti's (2020) survey, also shows that the current pandemic has increased buying online for the third of the Finns. This chapter's function is to guide the reader to the topic. The introduction briefly describes the background of the research, the research problems, the purpose and need of the research, the objectives, the research method and the results obtained and their significance. This thesis strives to provide a comprehensive vision of organization's data analytics capabilities by creating a typology based on the literature review's results. After that it is attempted to validate the factors that are relevant to e-commerce businesses. To achieve these objectives a literature review was executed to recognize the core capabilities related to data analytics in the context of e-commerce. Furthermore, ten semi-structured interviews were performed with professionals working in the e-commerce industry.

Interest towards data analytics has been rising exponentially, but the capabilities that organizations need to benefit from data analytics have not been considered enough in the context of e-commerce. To fill this gap, the following research questions were selected:

- What capabilities are required for companies to benefit from data analytics from the perspective of resource-based view?
- How are the different capabilities related to data analytics perceived in the case companies?

The first research question is answered on the basis of previous literature and the second research question is answered on the basis of empirical results.
The study ignores the use and utilization of artificial intelligence related to the topic to delimit the topic. The study is not intended to focus in detail on technological features, the introduction or implementation of data analytics tools.

Restrictions and regulations on consumer data, such as the General Data Protection Regulation (GDPR), which has entered into force in the European Union is not in the centre of this research, but it is mentioned because it appears in the interviewees' responses.

The theoretical part of the research consists of chapters 2 and 3. In the first theoretical chapter, chapter 2, the core concepts of the research are discussed which are transformed to create a conceptual basis for the theoretical framework of the research. This includes introducing the reader to the e-commerce and its characteristics, showing the connection to data analytics, describing the data's path from raw data to knowledge and its connection to the decision-making process and to the business value. Also, the categories of data analytics are introduced to create more clear understanding. The second theoretical chapter, chapter 3, seeks to answer the first actual research question that is: "Based on the current literature, what capabilities are required for companies to benefit from data analytics from the perspective of resource-based view?". To answer the question, a literature review was conducted, and a recourse-based perspective was selected to study the capabilities. The result of the literature review presents a comprehensive vision of organization's data analytics capabilities by creating a typology based on the relevant literature. Chapter 4 describes in more detail about the methodology of the empirical part. The methodology that was selected is a qualitative methodology and more specifically, ten semi-structured interviews were performed with professionals working in the industry of e-commerce. In addition, the chapter 4 explains in more detail how the empirical part and the interviews were constructed and analysed and presents the results of the research collected through the interviews. The results of the study are presented thematically.

The results are grouped under four main themes that are: 1) benefits of data analytics for e-commerce 2) identified problems of data analytics in e-commerce 3) data-driven decision making in e-commerce companies 4) capabilities of case companies and their evaluation. The subsection 4.2 describes the benefits of data analytics that appeared in the responses of the interviewees. The subsection 4.3 describes the problems that interviewees experienced related to data analytics in e-commerce. The subsection 4.4 describes data-driven decision-making in the case companies and how the interviewees experience data-driven decision-making in their own companies and, in the case of consulting companies, the level of data-driven decision-making in Finnish e-commerce according to their experiences. The final subsection 4.5 summarises interviewees' experiences under the resources corresponding to the results of the literature review and examines how well the results of the literature review match or differ from the interviewees' experiences. The final chapter, Chapter 5, presents the conclusions of the study and answers the second research question: "How are the different capabilities related to data analytics perceived in the case companies?" and compares the respondents' experiences with the results of the literature review, creating a comprehensive description of the current state of analytics in Finnish e-commerce companies. Finally, the success and limitations of the study are assessed and possible areas for further research are discussed.

2 CONCEPTUAL BASIS – DATA ANALYTICS AND DATA-DRIVEN DECISION-MAKING IN E-COMMERCE

This chapter introduces the current status and characteristics of e-commerce, shows the connection to data analytics, describes the data's path from raw data to knowledge and its connection to the decision-making process and to the business value. Also, the categories of data analytics are introduced to create a clear understanding of the topic.

2.1 E-commerce and its characteristics

Online shopping has become very common in the last decade and the trend seems to be that the share of online shopping will only grow in the future. E-commerce is generally perceived as a business where buyers purchase products, services or content online via computer networks. In e-commerce, customers are traditionally categorized into two: business-to-business (B2B) and business-to-consumer (B2C). In the literature, Maity and Dass (2014), define e-commerce-focused companies to be companies that sell services and products in an online platform. Other aspects that are strongly related to e-commerce that occur in the literature are, technology driven business processes and customer service (Kalakota & Whinston, 1997).

Newer definitions extend the definition to also include more customer-oriented themes such as digital value creation highlight e-commerce companies to be very customer-oriented organizations (Maity & Dass, 2014; Frost & Strauss, 2013). They also focus more on e-marketing and its importance. In this thesis, however, we are using even extended definition and try to include all the functions in e-commerce business that can benefit from data analytics. These functions include transaction values such as cost savings and efficiency improvements, inventory management and improvements in sales and marketing that are benefitted from data analytics. This same kind of wide definition that focuses on all functions that can be benefitted from data analytics have been also used by Akter and Wamba (2016) and Wixom, Yen and Relich (2013).

As the share of e-commerce have been growing, also the consumer behavior in e-commerce context have been studied widely. Studies have shown that when comparing e-commerce to traditional stores, trust and confidence rises as a priority (Chen & Dhillon, 2003). Other important features that customers appreciate are the ease of shopping, general functionality of the website, ease of returns, selection and freedom to make purchases at any time of the day (Posti, 2020). The ease of shopping includes in the case of e-commerce, that the customer doesn't have to physically go anywhere and this way the customer avoids queuing (Hyvönen & Pylvänäinen, 1999). Lampikoski and Lampikoski (2000) have also defined the freedom of shopping to mean that the shopping is not tied to stores opening hours, so the shopping can be done at any time that suits best for the customer. Many online stores provide free returns. In Finland, Consumer Protection Act ensures that all online purchases have 14-days return and exchange policy. This kind of policy protects the customer and also encourage the customer to buy products without physically seeing them. Most of the technically related features are strongly related to general aspects that effect on user acceptance of information technology (Davis, 1989). This is not surprising since these features are the same worldwide.

2.1.1 Reasons for the growth of e-commerce

As stated earlier, the popularity of e-commerce has been steadily growing. There are multiple causes to that. In addition to the increasing number of consumers, other side is that setting up an online store is easier than ever. There are many different e-commerce platforms that enables the merchant to open an online store without a single line of code.

Based on a survey, conducted by Paytrail (2020), the most popular e-commerce platforms in Finland are WooCommerce, MyCashflow, Clover Shop and Magento. Paytrail is a company currently owned by Nets Holding A/S and it offers payment services across the Nordic countries. The results of the survey are based on 508 responds from online retailers. According to BuiltWith (2021) WooCommerce is also the most popular worldwide but on the second place there is Shopify that has gained popularity during the last few years. Shopify is placed on fifth in Paytrail's survey, and its market share has increased from 2019.

These e-commerce platforms usually offer easy "drag and drop" website builders that doesn't require any coding or website developing skills. In addition to that, the platforms offer other additional features such as order management, inventory management, ready payment functions and marketing integrations to help merchants' tasks. They also offer reports about sales and inventory and it is easy to export data from the platform to analysis tools.

Besides the data from the platforms, also social media platforms, Google's marketing tool, Google Ads together with Google Analytics that tracks and reports website traffic, provide a lot of data to e-commerce merchants. They also provide ready information about merchant's marketing campaign's such as conversion rate and Click Through Rate (CTR) that especially people who work in marketing are monitoring. Social media marketing and search engine optimization are their own areas. They are not are not under the review of this thesis and according to Su et al. (2014) the validation of SEO's operating model is proved to be challenging in practice. The purpose of presenting briefly some of the social media tools is to bring up that they are one source of data for e-commerce companies.

As pointed out, the easy access and the availability of data together with analytical tools have enabled widespread of use of data analytics regardless of the size of the organizations. Besides the development of the analytical methods and tools, the technology-related costs have dropped dramatically (Acito & Khatri, 2014).

2.1.2 E-commerce data types

As discussed in the previous subsection, e-commerce merchants obtain data from a variety of sources and social media together with e-commerce platforms provide an easy access to the data. In ecommerce, data can be seen as the key to track consumer shopping behaviour to gather important information about the customer.

E-commerce businesses generally process data that can roughly be classified based on its structure, to unstructured data and structured data.

Unstructured data refers to data that doesn't have a predefined data model or doesn't fit into relational database tables (Sint et al., 2009). Typically interpreting unstructured data, such as videos, voices and photos, is easy for humans but for technological applications it can be arduous. Modifying unstructured data to fit data analytics often requires complex mathematical and statistical methods. In the context of e-commerce, unstructured data plays a big role because it also includes all the click-stream data that social media platforms provide (Akter & Wamba, 2016). This includes tweets, links, likes and clicks on social media content.

Structured data is characterized by having a clear structure and format and it can be analysed as such. All the customer information, such as demographic information, customer's name, age and address are examples of structured information in the context of e-commerce. Structured data in e-commerce includes all the transaction and business activity data (Akter & Wamba, 2016). E-commerce companies gather a lot of information over time while tracking consumer's browsing information.

Even though structured data can be seen as more critical for e-commerce businesses, it is important to notice that only a very small part of the data is in structured form (Gandomi & Haider, 2015, 138). Also, in the last decade, the amount of unstructured data has been growing rapidly (Khan et al., 2014).

2.2 Data analytics and e-commerce

In the last decade the interest towards data analytics has occurred in both scientific research and e-commerce industry. The interest in the topic has also been boosted by the trend related to big data and the hype surrounding it. Although the research area of data analytics is widely studied, there is no unambiguous definition of the concept of data analytics in the existing literature. Next the concept is defined to demonstrate its scope.

Some of the definitions focus mainly on the statistical and analytical methods to analyse data sets (Kwon et al., 2014; Müller et al., 2016), but most of the definitions also highlight that the focus for data analytics is aiding the decision-making (Davenport, 2006; Davenport & Harris, 2007). Other definitions also include improving the company's performance as well as optimizing business processes (Ghasemaghaei et al.,2015; Manyika et al., 2011; Kwon et al., 2014). To summarize the concept, it can be said that data analytics includes all the tools, methods, technologies related to data analysis that aims to generate valuable insight for the organization to improve the organization's performance.

The companies operating in the e-commerce industry are also said to be one of the fastest operators to adopt data analytics because they operate in a constantly changing business environment (Koirala, 2012). This kind of business environment forces e-commerce companies to maintain their competitiveness and continually seek for new ways to improve their business models and to find new business opportunities (Koirala, 2012). Data analytics have been proved to be a vital enabler of the improvements (Behl et al., 2019).

Due to the interest, all modern e-commerce organizations have started collecting enormous amounts of data from various sources. Organizations use analytical methods to create valuable insight and information to support organizations decision-making process and to gain competitive advantage (Akter & Wamba, 2016; Sumbal, Tsui, & See-to, 2017; Provost & Fawcett, 2013). Other identified incentives are the availability of data and the amount of the available data that is continually increasing (Provost & Fawcett, 2013).

2.2.1 From data to information

While researching data analytics and its connection to decision-making, it is important to briefly describe the path from data to wisdom. DIKW (data-information-knowledge-wisdom) hierarchy, which for example Rowley (2006) and others have researched, is a traditional way of describing this path. This hierarchy is often described as a pyramid (Figure 1).

Data is located at the bottom block of the pyramid and it reflects to the fact that it is the most available form of information. The higher we proceed in the pyramid the smaller the block gets. At the same time the block gets smaller, the more meaning and more value of the form of information has. For the sake of clarity, wisdom exists the least in the world, but it has the most value and meaning. According to this hierarchy, data consists of separate facts and chains of events without context, and they are products of observation. After data gets context, it turns into information. Information can answer traditional and simple questions such as how many, who, what and when. Information transforms into knowledge through assumptions and personal experience. The most significant difference between knowledge and wisdom is that with wisdom one can increase the effectiveness and wisdom adds more value to the decision-making process than knowledge (Rowley, 2006).

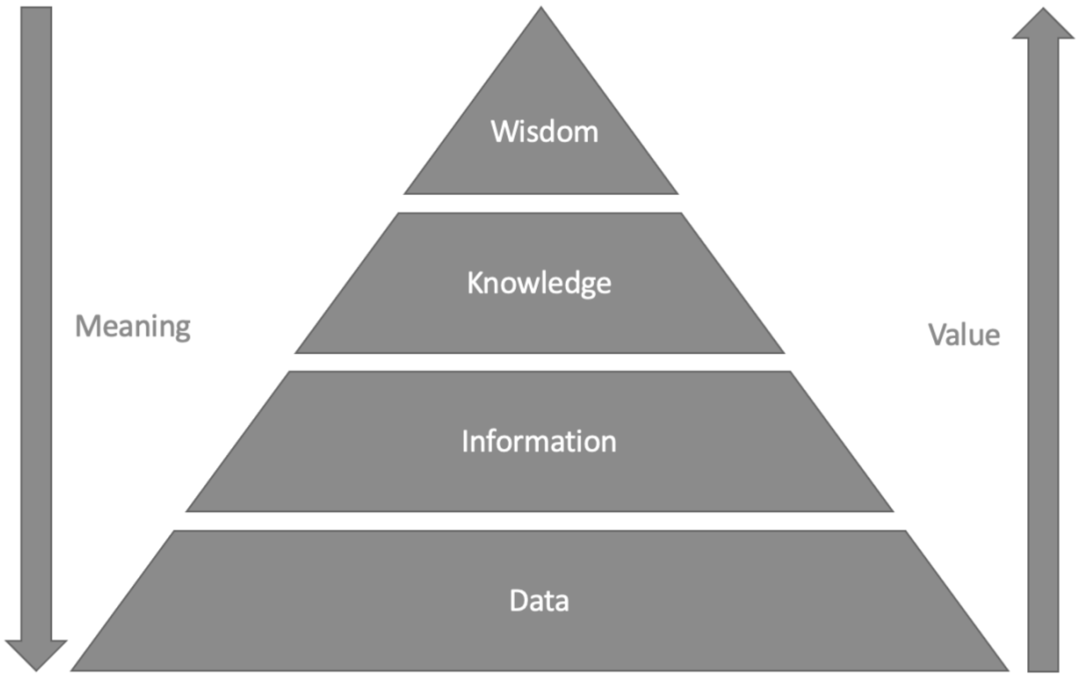

Figure 1 DIKW (data-information-knowledge-wisdom) hierarchy

Data science and data analytics tools are made to help in this process when we want to proceed higher in the pyramid. With these tools it is easier to collect and control data and to form it to information. There are different types of data analytics that are briefly introduced in next paragraph.

Wang et al. (2016) stated that data analytics can be roughly classified into three categories, which are: descriptive analytics, predictive analytics, and prescriptive analytics. Descriptive analytics is one of the most known types of data analytics. Descriptive analytics refers to analyses that are performed regularly or as needed and are designed to identify problems as well as opportunities based on the basic existing data (Wang et al., 2016). Descriptive analytics is used to find trends in data or repetitive patterns of behavior. For example, it is possible to monitor purchasing behavior and descriptive analytics can help organizations to create customer segments to target the desired target customer segment (Davenport, 2006). In practice, descriptive analytics seeks to answer the questions, “What happened?” and “What’s going on?”.

According to Demirka and Delein (2013), predictive analytics differs from descriptive analytics in that way that it uses mathematical models and algorithms to find explanatory and predictive models. In predictive analytics, historical data is analysed to detect recurring patterns and trends so data can be turned into predictions. The purpose of predictive analytics is to answer the question, "What will happen in the future?" as well as justify why this can happen. However, Lawless (2014) emphasizes that the purpose of predictive analytics is not to tell with certainty what will happen next, but to analyse possible events with moderate reliability. Key tools that enable predictive analytics includes more data mining and more data types than only text data (Demirkan & Delen, 2013). Practical example of the use of such analytics in e-commerce is a situation where predictive analytics make it possible to track demand trends or tell how a marketing campaign affects the customers consuming behavior of a certain customer segment (Raden, 2010).

Prescriptive analytics includes the use of data, mathematical models and algorithms to create, define, and evaluate alternative options. Characteristics for the models are large data amounts as well as complexity. Prescriptive analytics also includes multi-criteria decision-making, optimization, and simulation (Souza, 2014). The purpose of prescriptive analytics is to make recommendations for improving the business by answering the question "What should happen to get the business at a certain level?" (Souza, 2014).

2.2.2 Data-driven decision-making

Different types of data analytics were described above and the role of data analytics in supporting decision-making was mentioned. Decision-making can be seen as the most significant factor influencing organization's performance because decisions constantly control and guide all the activities of the organization. Decision-making is also an integral part of the process where data analytics is utilized. According to Porter & Advantage (1985) the organizations failure or success depends primarily on the supervisors' ability to make decisions in a competitive business environment. It is therefore natural to view more deeply at decision-making.

There are multiple different angles for viewing organizational decision-making. One widely accepted angle is to categorize decisions to structured decisions and unstructured decisions (Scherpereel, 2006). In this angle the decisions are categorized based on the complexity of the subject matter (Turban, Aronson & Liang, 2005). Depending on the complexity, the organizational decision-making process can be unstructured or structured (Langley et al., 1995).

Unstructured decisions refer to a decision-making process that hasn't been made before in an organization, so there are no readily available answers to it (Mintzberg, 1978). Because of that, unstructured decisions are usually based on manager's previous experience and intuition. Data-driven decision-making (DDD) refers to the activities where decisions are made based on the analysis of data rather than based on intuition (Brynjolfsson, Hitt & Kim, 2011; Provost & Fawcett, 2013). The proponents of DDD disagree on that the unstructured decisions are based on previous experience and intuition and they see data as one core element among experience and intuition.

Traditionally, structured decisions can be described by using classic mathematical models, while unstructured decision don't have standardized methods for obtaining the most optimal solution (Zhang, Lu, Gao, 2015). Drucker (1967) has developed systematic process for structured decision-making. In that process, the problem is firstly classified and secondly the problem is defined. Thirdly, the wanted outcome is defined. Fourthly, the conditions are defined. Lastly, the implication plan of the decision is defined and the validity and effectiveness of the decision for the problem is tested. Structured decisions are made in day-to-day operations to achieve short term goals. This type of decision-making is usually used for routine decision-making. However, decision-making is not always a linear process, and it can be seen rather as a dynamic, cyclical process in a complex business environment that is also influenced by the interaction of people. Linear and complex decision-making cannot be completely separated, and it would be unlikely that an organization would only use one kind of decision-making process in their decision-making.

Heisig et al. (2016) determine the future of information management to focus on activities and processes that promote the creation of individual and organizational level knowledge resources to achieve competitive advantage. They identify the most important future research topics to be intangible capital, knowledge sharing, organizational level learning, innovation and achieving the competitive advantage.

Like stated earlier, data-driven decision-making refers to the activities where decisions are made based on the analysis of data rather than only based on intuition (Provost & Fawcett, 2013). It is also important to notice that in data-driven decision-making the decisions don't need to rely only on the information from data analysis, but it can be a combination of intuition and knowledge based on the data analysis (Provost & Fawcett, 2013). Data driven decision-making is also often incorporated with organizations performance (Ghasemaghaei et al. (2015). Improved performance is seen to be due to process optimization (Shanks et al. (2010). The connection between data and decision-making has been illustrated by Intezari and Gressel (2017) in their data-decision quadrants (Figure 2).

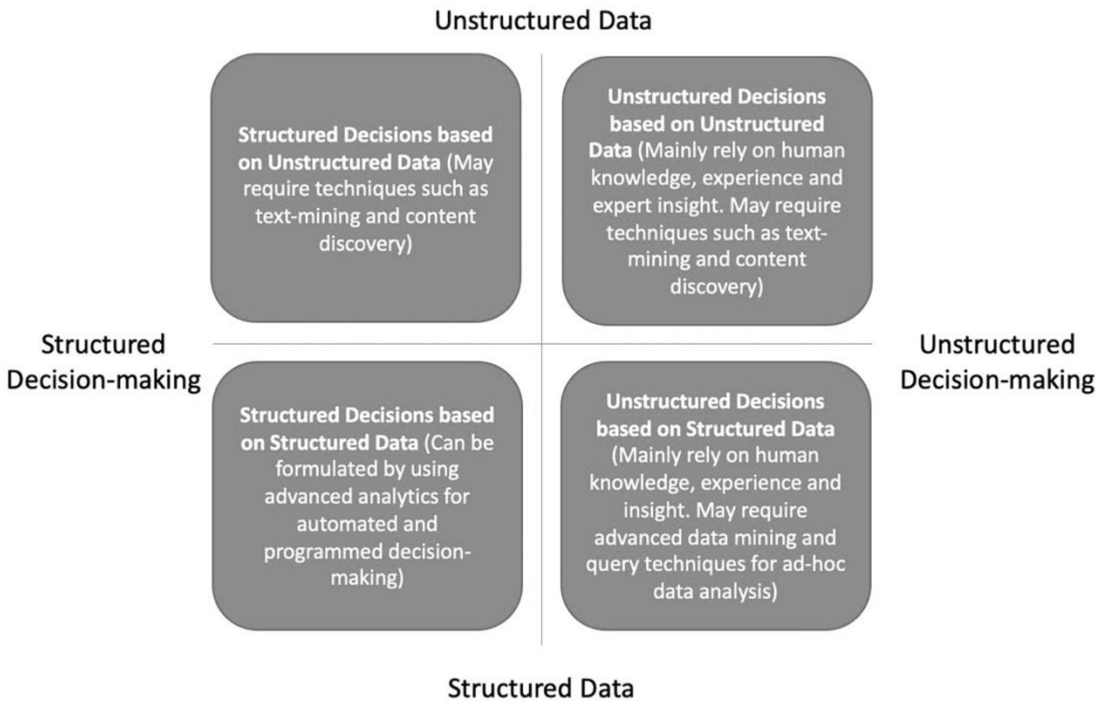

Figure 2 The decision-data quadrants (Intezari & Gressel, 2017)

The quadrants main idea is to illustrate that there are four major types of decision-making. The type of decision-making depends on the type of data and the type of decision-making. The main finding of this figure is that unstructured decision-making should be almost always be done based on data (Intezari & Gressel, 2017). It is also noteworthy that these decision-making models can be used crosswise depending on the situation and they are not mutually exclusive. Organization should also adapt to the situation and the type of selected decision-making depends on the availability and quality of data (Intezari & Gressel, 2017).

Provost and Fawcett (2013) acknowledge in their research that there is a undeniable proof that the data-driven decision-making with the help from data science have beneficial effects on organizations performance and competitive advantage. In the existing literature the benefits of data analytics are incorporated either to decision-making process or to company’s performance. In the case of decision-making, the analysis is seen as an aiding factor in the decision-making process that facilitates the management’s actions and decisions (Davenport, 2006; Davenport & Harris, 2007). The same subject has been studied by others who also confirm that the connection between data-driven decision-making and data-science techniques. They also confirm that data-driven decision-making improves organization’s performance and the ability to gain business value from data (Brynjolfsson, Hitt & Kim, 2011; Hill, Provost & Volinsky, 2006; Martens & Provost, 2011).

Although the benefits of data-driven decision-making are clear, studies have also identified situations where the use of data analytics does not benefit decision-making. According to Davenport, Harris and Morison (2010) there are five types of these situations. The first one is a situation where there is no time to do the analysis. Such a situation, where time is scarce and the analysis is done in a hurry, can lead to distorted results, which only reduces the possibility of making the right decisions. The second one is a situation where there is no previous knowledge available. In this kind of situation, the use of data analytics may again produce distorted information if the data that is used in the analysis does not reflect the situation or is inappropriate for the situation. Third situation that Davenport et al. (2010) list is a situation where the history is misleading. Misleading history refers to a situation that exploits precedents whose variables are not fully known. Such an analysis cannot be relied upon due to the missing factors. The fourth situation is not as strongly related to data or analytical tools as the previous ones. The fourth situation is a situation where the decision maker has a proper amount of experience already to make the decision. This is rationalized by the fact that repeating the process can be pointless and a waste of time. The final situation that was listed is a situation where the variables are not measurable. This refers to situations where key variables cannot be reliably measured and converted into analytical factors. If the key variables are forcibly converted to formats that can be analysed, it can lead to incorrect information which hampers the decision-making (Davenport et al., 2010).

Other interesting point of view is presented by Lucker and Guszcza (2012), who compiled a list of the five most common errors associated with the use of analytics in business intelligence analytics. Unlike Davenport et al. (2012), the reasons they list, are not due to the inappropriateness of the analytics, but are due to people's expectations related to the analysis that differ from reality. Lucker and Guszcza both worked for Deloitte Consulting LLP at the time of publication. The first one in their list is the misunderstanding of the analytics. By this they mean that data analytics should not be considered as a mysterious forecasting tool that can predict the future. They also state that managers tend to fail to understand what data analytics can do in practice. The second general error is that managers concentrate too much on construction of databases before they try to benefit from data. The third one is that data analyst that is responsible for the data tries to achieve mathematical absolute truths rather than efficiently and quickly utilizing the processed data. The fourth error is that some managers rely too blindly on the results that data analytics provide without the necessary critique and at the same time they fail to assess the relevance of the results. The last one is a situation where there is a lack of communication between data experts and decision makers. In such a situation, the benefits of data analytics often remain very small.

3 LITERATURE REVIEW – DATA ANALYTICS CAPABILITIES

As previously shown, the benefits of data analytics are undeniable for e-commerce companies and for their decision-making. In this chapter the concept of data analytics capabilities is described and the resource-based view, that is strictly connected to this topic, is discussed. After that, a narrative literature review was conducted to create a typology that provides a comprehensive vision of organizations' data analytics capabilities and to recognize the core capabilities in the e-commerce industry.

Literature review's function in general is to act as a link between a vast amount of literature and the author that isn't able to analyse all the literature related to the topic (Baumeister & Leary, 1997). Narrative literature review is described as an overview without strict and precise rules, and it aims to transform incoherent information to an easily understandable entirety (Freeman, 1984). The benefits of narrative literature review are that the examined phenomenon can be described broadly, and the characteristics of the phenomenon can be classified easily (Freeman, 1984). Based on this, a narrative literature review supports our needs to create a comprehensive synthesis of previously published literature. According to Webster and Watson (2002), many authors fail to synthesize the literature by doing just concept-centric or author-centric literature review. In this literature review, a concept matrix is used to avoid that failure and to identify the key concepts and characteristics of this topic.

The source material for the literature review was collected systematically by using scientific databases. The chosen databases are Google Scholar, Science Direct and IEEE Xplore. For the actual search an advanced search feature was used. Advanced search enables the use of different combinations of keywords. The source material was compiled by using the following keyword combinations:

data analytics capabilities AND e-commerce

data analytics AND capabilities / capability

The first query was selected because it binds the two key concepts in this topic together. The second query was selected to bring more characteristics related to data analytics capabilities outside the e-commerce industry. This enables us to possibly find new characteristics outside the industry so we can make new findings. It also supports the main purpose of this literature review by creating a comprehensive view to this topic.

From the Google Scholar search alone, the first query produced around 72 200 results and the second query produced around 137 000 results. The subject has been studied extensively and in order to find appropriate studies from the existing literature, the results were limited based on their relevance and year of publication. The year of publication was limited to the last fifteen years, but the aim was to emphasize as many new studies as possible in the results. Also, the results were selected mainly from studies focusing on e-commerce. In addition to this, articles related to artificial intelligence were excluded from the literature review because they are excluded from the scope of this thesis. Nor is it intended to focus in detail on technological features or implementation of data analytics tools.

3.1 Theoretical background of analytics capabilities

3.1.1 Resource-based view

Before moving onto organization's data analytical capabilities, it is important to open the perspective that we are examining. The perspective that was selected is resource-based view (RBV) that was presented by Barney (1991). According to resource-based view, an organization's competitiveness is based on resources that are valuable, infrequent, difficult to imitate and difficult to replace.

Valuable resources allow an organization to increase its revenue and to reduce its costs. The second type, infrequent resources describe the resources that only a few companies have, and those resources increase the company's competitiveness. The resources that are difficult to imitate refer to resources that cannot be copied directly and others can't create exactly the same kind of resources. Expensiveness can be considered as one factor that makes the resource difficult to imitate. The recourses that are difficult to replace are usually considered as organizational resources that emphasizes the importance of management. To benefit from the previous resource types, it is important that the management can manage the other resources properly. Barney, Wright and Ketchen (2001) defined the resources and capabilities to be:

> "bundles of tangible and intangible assets, including a firm's management skills, its organizational processes and routines, and the information and knowledge it controls." (Barney, Wright & Ketchen, 2001)

Based on the RBV, also Grant (1996) presented his view that was more focused on the organization's knowledge capital. According to Grant's (1996) knowledge-based view, company's competitiveness is based on the organization's internal characteristics that are related to knowledge. In knowledge-based view the company's performance arise from differences in information resources and company's ability to utilize and develop new information (Grant, 1996). Knowledge can be seen as the most important resource to a modern organization and the ability to get leverage from knowledge resources is critical for any organization. From this a direct connection can be drawn to data-driven decision-making and to organizations learning culture which is reflected to knowledge management (Teece, 2014).

The Barney's resource-based theory has evolved into a general method to describe, explain and predict organizational connections in business economics and to identify the company's strategic resources (Barney, Ketchen & Wright, 2011; Kozlenkova, Samaha & Palmatier, 2014). Resource-based theory is directly related to data analytics as the exploitation of data seeks to use organizational resources to achieve competitive advantage. Several studies state that there is a lot of potential in resource-based theory to create an all-encompassing strategic theory for organization (Mahoney & Pandian, 1992; Palmatier, Dant & Grewal, 2007). Studies have also found that, while comparing resource-based theory to contingency theory, resource-based theory has a stronger predictive ability of IT impact on company's profitability and revenue (Oh & Pinsonneault, 2007).

In a recent study, Akter & Wamba (2016) draw a connection between RBV and data analytics in the field of e-commerce and data analytics. They argue that the use of data analytics is a distinctive competence that companies need to enable high-efficiency processes. These kinds of supporting benefits that support business processes can be for example identifying the customers with the best return in their life cycle, optimizing the price or predicting the lowest possible inventory level (Akter & Wamba, 2016; Davenport & Harris, 2017). Devaraj, Fan & Kohli (2002) have studied the same area from the transaction cost theory point of view that was presented by Williamson (1981). In their research they found out that organizations can leverage from data analytics by making the transactions cost more efficient (Devaraj, Fan & Kohli, 2002). The efficiency was enabled by data analytics because it saves time and can give recommendations to company's management.

3.1.2 Classification of data analytics capabilities

Data analytics can be used to manage large amounts of data. In this case with data analytics, we refer to data analytics where large amounts of data are transformed into information with tools such as data mining, visualization and statistical analysis (Chen, Chiang & Storey, 2012). Ghasemaghaei, Ebrahimi and Hassanein (2018) emphasizes in their study the importance of decision-making related to data analytics. They define the data analytics as a combination of processes and tools that retrieve valuable perspectives from large – and potentially fragmented – data sets to support organization's decision-making.

The data analytics are used to improve the organizations' decision-making and with that the aim is to improve the organization's performance and to gain more business value. However, studies show that the vast majority of organizations investing in data analytics have not benefited significantly from them (Ghasemaghaei, Ebrahimi and Hassanein, 2018). Poor efficiency may be due to poor data quality, inappropriate data analytic tools, or because of the lack of analytical skills. It is said that organizations can also fail to leverage the information that they obtained from data analytics (Ghasemaghaei, Ebrahimi and Hassanein, 2018; Ross, Beath & Quaadgras, 2013). Data analytics capabilities – the ability to utilize from data analytics – are essential to modern organizations and therefore, it is also necessary to study them more in practice to understand what the limiting factors in the context of e-commerce are.

According to Ghasemaghaei, Ebrahimi & Hassanein (2018) data analytics capabilities consists of five components. The components are data quality, data quantity, analytical skills, domain knowledge and sophistication of methods. All of these components have been found to have a significant positive effect on decision-making quality. All except data quantity has also been found to have significant effect on decision-making efficiency (Ghasemaghaei, Ebrahimi & Hassanein, 2018). Gupta & George (2016) also found data analytics capabilities to improve company's performance. They created a classification that is based on the RBV that was introduced earlier in this chapter. In this classification the data analytics capabilities are classified into three sections based on the resource type. These sections are tangible resources, human resources and intangible resources (Gupta & George, 2016). The same classification is used to compile the results of the literature review. Results of the literature review are presented in Tables 1-3. To illustrate the topic, a conceptual framework has been created, which is presented in the Figure 3 below.

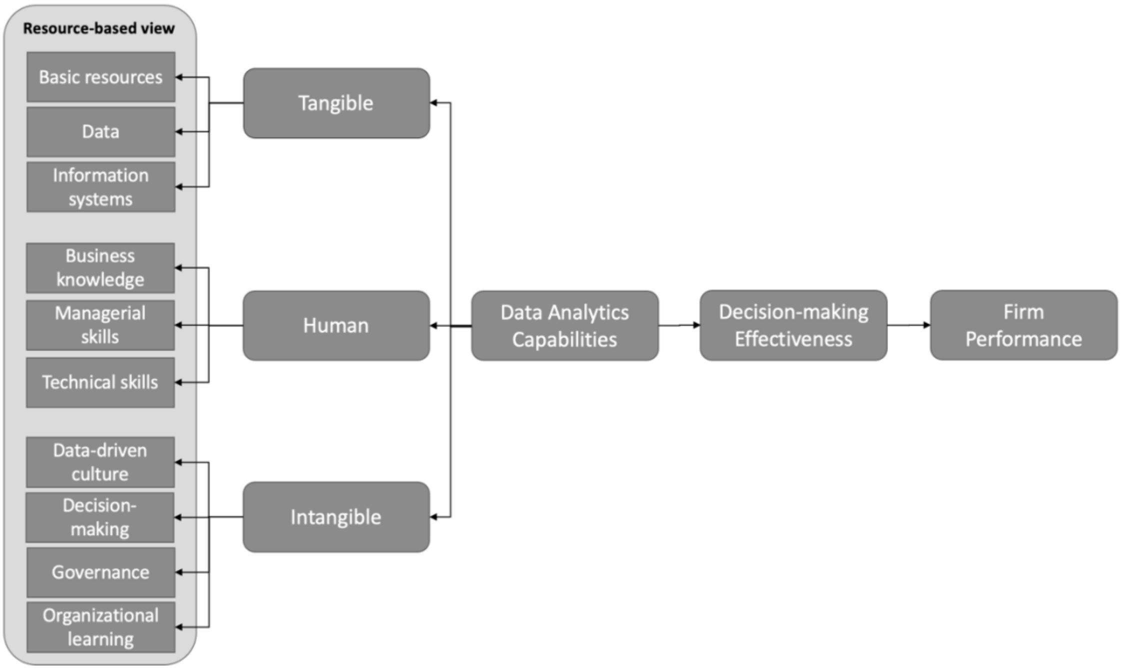

Figure 3 Conceptual framework

3.2 Identified tangible resources

A typical feature for tangible resources is that they can often be bought, and they are often physical such as data and different software. However, there are exceptions and also time is considered as a tangible resource (Gupta & George, 2016). Tangible resources can already exist in an organization, but tangible resources do not usually create competitive advantage or value on their own (Gupta & George, 2016). Tangible resources are still highly necessary for organizations while reviewing its data analytics capabilities.

The first identified tangible resources are basic resources. With basic resources, in this context, we refer to resources that are not directly related to data analytics. Basic resources are also slightly different than other identified resources because all organizations have these resources. Two basic resources appeared in the literature. The first one of the basic resources that occurred in the literature is time (Davenport, 2006: Gupta & George, 2016) and the second one is monetary investments (Gupta & George, 2016; Mikalef et al., 2020). Monetary investments are also highly connected to the decision-makers' willingness to invest in data analytics. Basic resources are important to mention in the context of data analytics capabilities because they are needed to create and to enable other resources which will be presented afterwards.

Data is a widely identified and also very critical resource because it is also a prerequisite for the data analysis. To benefit from data, the access to the data needs to be uncomplicated and the production of the data needs to be continuous (Carmichael et al., 2011). Organizations that own their data have better accessibility to the data than organizations where the data is processed by an external operator (Mithas et al., 2013). Traditional characteristic that is required from data is the sufficient quantity of the data (Ghasemaghaei et al., 2018; Davenport et al., 2012). Without that, it is impossible to create reliable analysis. In addition to the amount of data, a lot of attention should also be paid to the quality of the data (Ghasemaghaei et al., 2018). To create accurate analysis and information, the data needs to be accurate because if the data isn't accurate, the analysis can produce false information and thus leads to erroneous conclusions (Chae et al., 2014). In addition to the accuracy, the data needs to also be relevant and be applicable for that specific situation (Barret et al., 2015). The last identified characteristic is merging internal and external data. It is often difficult for many organizations. This characteristic is not directly related to data alone and it requires also technical and other skills to be able to merge and manage internal and external data to create extensive and comprehensive data sets for analysis (Gupta & George, 2016; Mikalef et al., 2020).

Third tangible resource was named as information systems. This resource is more multidimensional than the previous ones and for clarity it was justified that all the resources that are related to information systems were placed under this concept (e.g., platform, technology and infrastructure) to clarify the compilation of results. Xiao et al. (2020) talks about information systems as a resource, but Kiron et al. (2012) bring up the importance of the platform and Mithas et al. (2011) focus more on the IT infrastructure. The platform and the whole IT infrastructure are necessary because the adoption of data analytics requires technically sound infrastructure (Behl et al., 2019). Some authors talk about technological capabilities that are essential to explore and to manage different types of data (Barton & Court, 2012). The unifying factor, however, is that whether talking about platform or information systems in general, it is considered important that the systems are accurate, timely, reliable secure and confidential (Mikalef et al., 2020; Mithas et al., 2011). Security and privacy should also be mentioned at this point, as they relate to information systems, although they are not the subject of this thesis. Secure systems ensure the secure and safe processing of data which can be seen as very important for e-commerce companies as they often handle sensitive customer data. (Akter & Wamba, 2016; Mithas et al., 2011). Results related to tangible resources are presented in Table 1. The first column of the table indicates the type of the identified resource. The second column of the table indicates the identified resources and resources' typical characteristics that were found are listed below each resource. The third column of the table indicates authors whose papers were used to identify each resource. Same columns are used also in Table 2 and Table 3.

TABLE 1 Results of the literature review related to tangible resources

TYPE OF THE CAPABILITY	**Identified resources and *characteristics***	**Author(s)**

TANGIBLE	Basic resources: • *Investments* • *Time*	Davenport, 2006; Gupta & George, 2016; Mikalef et al., 2020
	Data: • *Access to data* • *Accuracy, data needs to be accurate so organization can gain value for analyzing it* • *Amount of data* • *Data quality* • *Data quantity* • *Merging internal and external data* • *Velocity and variety of data*	Barrett et al., 2015; Chae et al., 2014; Davenport, Barth & Bean, 2012; Ghasemaghaei, Ebrahimi & Hassanein, 2018; Gupta & George, 2016; Mikalef et al., 2020; Mithas et al., 2013; Carmichael, Palacios-Marques & Gil-Pechuan, 2011
	Information systems (Platform, Technology and Infrastructure): • *Accuracy, timeliness, reliability, security and confidentiality of the systems* • *Technological capabilities are essential for exploring and managing variety of data*	Akter & Wamba, 2016; Barton & Court, 2012; Behl et al., 2019; Kiron, Prentice & Ferguson, 2012; Mikalef et al., 2020; Mithas, Ramasubbu & Sambamurthy, 2011; Xiao, Tian & Mao, 2020

3.3 Identified human resources

Human resources include all the skills and capabilities that a employees' have. These can relate to employees' knowledge, technical skills, experiences, management skills, communication skills or other abilities that an employee has (Barney, 1991). In the context of data analytics capabilities, the technical and managerial skills are is said to be the most critical ones of human resources (Gupta & George, 2016). Three different resources were identified from the existing literature in the literature review. These resources are business knowledge and business analytics, managerial skills and technical knowledge and technical skills.

Business knowledge and business analytics was classified as its own resource, although this resource could have been classified to technical skills and management skills. This is because the business knowledge that includes the knowledge of the field isn't really a technical or managerial skill that one can learn but it is something that one achieves through their experience and by following the business environment's changes (Ghasemaghaei et al., 2018). The better vision the organization has of the company's operations and its business environment, the better organization can utilize the information provided by data analytics. Without the clear understanding of the data, it is impossible for the organization to benefit from the data (Chae et al., 2014). Business knowledge is often compound to business analytics and analytic skills in the literature (Kiron et al., 2012). Business knowledge and business analytics is also associated with organization's performance and business strategy alignment (Akter et al., 2016).

Managerial skills refer to the ability of management to utilize data analytics in decision-making. This includes managers' business acumen as well as understanding of the data and how to use it (Gupta & George, 2016; Davenport et al., 2012). Optimizing of the decision-making processes is attached to the same context and it means that the management should have the ability to optimize decision-making models so the decision-making processes could be faster, more accurate and more efficient (Barton & Court, 2012). Also, managers' attitudes toward data analytics are seen as a critical factor in this resource (Behl et al., 2019). If managers don't find data analytics useful, they will not incorporate it into decision-making. In that case, the benefits of data analytics often remain small and other resources invested in analytics work are wasted. In the literature, leadership is seen as an enabler of change and managers should have IT skills or at least a good understanding what kind of information data analytics can provide, and they should know how to request and consume data analyses (Manyika et al., 2011; Chandrasekaran et al., 2013). Akter and Wamba (2016) bring out that analytics driven management culture helps in the utilization of data analytics and the general interest towards data analytics in management also enables better decision-making and better utilization of data analytics. Data-driven culture is its own resource, and it is one of the intangible resources, but the analytics driven management culture can be separated into its own sub-area, which is why it is already presented at this point. The key to the formation of analytics driven management culture is the recognition of the value of data analytics and the constant increase in awareness (Mikalef et al., 2020). Brinkhues et al. (2014) also argue that managerial skills are the only potential source for achieving a sustainable competitive advantage. In short, it can be said that without managers' will and belief in the value of data analytics, the full value of data analytics will not be achieved.

Technical knowledge and technical skills of the employees are a natural resource for an organization that wants to benefit from data analytics. This resource has been identified in several different studies and this resource contains more than just coding skills (Davenport et al., 2012; Xiao et al., 2020). Technical knowledge and technical skills are categorized into human resources. This resource has also been identified to be a key resource for gaining competitive advantage (Mikalef et al., 2020). Technical understanding and skills enable the organization to reach full potential of the data analytics and they act as a critical part of utilizing data analytics (Vossen, 2014). The use of sophisticated analytical methods is also linked to the same context because the use requires a deep technical understanding (Ghasemaghaei et al., 2018). According to Lamba and Dubey (2015), many organizations aren't able to benefit from data analytics because they don't have the right talents. Data analytics is said to also require strategists and tacticians, in addition to skillful data scientists, who understand how to organize integrate data analytics as part of business processes (Lamba & Dubey, 2015). A potential challenge in this resource is seen to be that it may be difficult for employees to maintain their skills and knowledge as data analytics and tools evolve (Gandomi & Haider, 2015). However, data analytics has recently become quite trendy and interest towards data analytics has grown (Behl et al., 2019; Gandomi & Haider, 2015). The number of different online trainings and university courses has also increased, which will hopefully meet the need related to the number of experts. Results related to human resources are presented in Table 2.

TABLE 2 Results of the literature review related to human resources

TYPE OF THE CAPABILITY	Identified resources and *characteristics*	Author(s)
HUMAN	Business knowledge and business analytics: • *Knowledge of the field* • *Organizations need to have clear understanding of the data and analytics to benefit from them*	Akter et al., 2016; Chae et al., 2014; Ghasemaghaei et al., 2018; Kiron et al., 2012
	Managerial skills: • *Leadership seen as an enabler of change* • *Management's capability to optimize decision-making models* • *Managerial IT skills to gain sustainable competitive advantage* • *Manager's ability to request and consume data analyses* • *Manager's business acumen and understanding of the data and how to use it*	Akter & Wamba, 2016; Barton & Court, 2012; Behl et al., 2019; Brinkhues, Maçada & Casalinho, 2014; Chandrasekaran et al., 2013; Davenport et al., 2012; Gupta & George, 2016; Manyika et al., 2011; Mikalef et al., 2020; Teece, 2014
	Technical knowledge and technical skills: • *Coding skills* • *Skilled data scientists* • *Sophisticated methods* • *Strategists who understand how to deploy the tool* • *Supporting technology personnel to implement the data* • *Tacticians to organize and manipulate data into operational models*	Behl et al., 2019; Davenport et al., 2012; Ghasemaghaei et al., 2018; Gandomi & Haider, 2015; Lamba & Dubey, 2015; Manyika et al., 2011; Mikalef et al., 2020; Vossen, 2014; Xiao et al., 2020

3.4 Identified intangible resources

The specification of intangible resources isn't as unambiguous as the previous ones because intangible resources are rather abstract. Intangible resources are generally not available for purchase except trademarks and copyrights. A good example of intangible resources is an organization's culture. Typically, intangible resources don't have clear boundaries and they are not heterogeneous across the organization and they are difficult to transfer (Teece, 2014). Despite the difficult nature of the intangible resources, Gupta & George (2016) identified that in the context of data analytics, the most important intangible resources are the intensity of learning in the organization and the culture of the data-driven decision-making.

Data-driven culture is a resource that is hard to measure, and it consists of employee attitudes, processes, utilization of data analytics, and development of new innovations. Many organizations want to achieve data-related cultural change, but especially for traditional and large organizations, this can be really hard because practices and processes are so established. Analytics driven management culture that Akter and Wamba (2016) introduces was mentioned already earlier in the context of managerial skills but in this section, we look at it more holistically. Data-driven culture is a critical resource especially when an organization is deploying new data analytics tools or tries to change existing practices related to data analytics (Mikalef et al., 2020). Data-driven culture is also a key factor when determining data analytic projects success and continuation (LaValle et al., 2011). Organizations that have a strong data-driven culture are described as organizations that use data in a pervasive way and develop processes to make it easy for employees to acquire information. In these organizations the decisions are made based on data rather than intuition and employees have a desire to seek how to benefit from data analytics (Gupta & George, 2016). Organizations that have strong data-driven culture can also be seen as rapidly adaptable and results-oriented organizations where the managers are able to "break out of their box" and align all eyes on the target (Lamba & Dubey, 2015). Data-driven culture also creates processes around data analytics which can create new innovations and new ideas how to benefit from data (Mikalef et al., 2019).

According to Gupta & George (2016), decisions in modern organizations should be make based on data rather than intuition. Decision-making is considered an intangible resource, even though the actions resulting from the decisions are tangible. Decision-making and data analytics go hand in hand and as noted earlier, data analytics is useless if the information that it provides does not lead to any action. That is why decision-making as a resource can be seen extremely important because without it, the organization is just wastes other resources that is use for data analytics. The organization shouldn't therefore see data analytics to have intrinsic value but rather have instrumental value.

Benefitting from data analytics requires quick response and timely decision-making processes (Xiao, Tian & Mao, 2020). Decision-making is also influenced by decision-makers desire to seek how to benefit from data analytics and this desire is seen to have a very positive influence on using the information that data analytics provides in decision-making processes (Gupta & George, 2016). This resource is also influenced by other resources such as managerial skills, business understanding and the prevailing culture in the company. Lamba & Dubey (2015) also bring up the importance of the strategy meaning that organization should have the use of data analytics in their strategy. In data-driven decision making the data validates the decisions and provides certainty for the decision makers.

Governance is an intangible resource that covers data management, practices, standards and rules for using data in an organization. Successful governance supports data-driven culture and makes the data access restrictions and governance practices transparent throughout the organization (Tallon, Ramirez & Short, 2013). At its best, proper governance is effective and the ownership of the data is clear, and it also improves the accessibility of the data (LaValle et al., 2011).

There are multiple practices that help an organization to construct better governance. LaValle et al. (2011) has presented a few of them. First practice is advancing standard methods for identifying business problems to be solved with analytics. Second practice is facilitating identification of analytic business needs while driving rigor into methods for embedding insights into end-to-end processes. Third practice is promoting enterprise-level governance on prioritization, master data sources and reuse to capture enterprise efficiencies and the last practice is standardizing tools and analytic platforms to enable sharing, streamline maintenance and reduce licensing expenses. The right kind and functional governance also ensure trust and enables secure and lawful analysis of data in the organization (Davenport et al., 2007).

Organizational learning is an intangible resource that is connected with innovative culture. Organizational learning means the active development of the skills and increasing the knowledge of employees so that data analytics related knowledge and skills are not left to data analysts alone. In practice this resource appears as an organizational ability to train required skills when needed and it is also important that the organization encourages the employees to share knowledge between each other (Gupta & George, 2016).

Various online courses and organizations' training portals have facilitated the acquisition and maintenance of this resource and it has been shown that further training of existing employees is often more beneficial and more cost-effective than recruiting new ones (Kiron, Prentice, & Ferguson, 2012). However, learning takes place continuously without courses and it would be particularly important for the organization to be able to share information and enable learning from other employees. Innovative culture seen as a contributing factor to this (Mikalef et al., 2020). Ideally, employees themselves recognize the potential benefits of data analytics in their own work, in addition to which the company encourages to learning and innovating and the development of working methods. In this manner, the data analytics will be part of everyone's job, and it can be seen as an ongoing process regardless of the department the employee is working in (Davenport, 2006). Results related to intangible resources are presented in Table 3.

TABLE 3 Results of the literature review related to intangible resources

TYPE OF THE CAPABILITY	**Identified resources and *characteristics***	**Author(s)**
INTANGIBLE	Data-driven culture: • *Analytics driven management culture.* • *Desire to seek how to benefit from data analytics*	Akter & Wamba, 2016; Gupta & George, 2016; Lamba & Dubey, 2015; LaValle et al., 2011; Mikalef et al., 2020; Mikalef et al., 2019
	Decision-making: • *Decisions are made based on data rather than institution* • *Predefined strategy for using data analytics*	Gupta & George, 2016; Lamba & Dubey, 2015; Xiao et al., 2020
	Governance: • *Data management* • *Ensuring trust*	Davenport et al., 2007; LaValle et al., 2011; Tallon, Ramirez & Short, 2013
	Organizational learning and innovative culture: • *Ability to share knowledge* • *Ability to train required skills when needed* • *Data analytics as a part of everyone's job as an ongoing process* • *Learning about data analytics across the organization*	Davenport, 2006; Gupta & George, 2016; Kiron et al., 2012; Mikalef et al., 2020

4 EMPIRICAL RESEARCH AND RESULTS

This chapter consists of two parts, the first one aims to describe the methodology of the empirical part and the second part presents the results of the research collected through the interviews. The methodology that was selected is a qualitative methodology and more specifically, several semi-structured interviews were performed with professionals working in the industry of e-commerce. These methods are described in more detail in the subsections of this chapter. In addition, the chapter explains in more detail how the empirical part and the interviews were constructed and analysed.

The results of the study are presented thematically, each in its own subsection, but some of the themes that are discussed are linked to each other, so the connection between the themes is also addressed. The thematic areas are derived from the theoretical basis of the research and the most important topics are highlighted by direct quotations from the interviews. The results are grouped under four main themes that are: 1) benefits of data analytics for e-commerce 2) identified problems of data analytics in e-commerce 3) data-driven decision making in e-commerce companies 4) capabilities of case companies and their evaluation.

The subsection 4.2 describes the benefits of data analytics that appeared in the responses of the interviewees. The subsection 4.3 describes the problems that interviewees experienced related to data analytics in e-commerce. The subsection 4.4 describes data-driven decision-making in the case companies and how the interviewees experience data-driven decision-making in their own companies and, in the case of consulting companies, the level of data-driven decision-making in Finnish e-commerce according to their experiences. The final subsection 4.5 summarises interviewees' experiences under the resources corresponding to the results of the literature review and examines how well the results of the literature review match or differ from the interviewees' experiences.

4.1 Methodology

According to Hirsjärvi et al. (2009), a research strategy refers to a decision of how the research will be conducted and research method refers to the technique by which the research is executed. The choice of research strategy as well as the choice of single research methods depends on the chosen research problem (Hirsjärvi et al., 2009). Research strategies can be categorized into three main groups: experimental research, quantitative research and qualitative research. In qualitative research the aim is to seek detailed information about small group of people that are in certain ways connected to each other (Hirsjärvi et al., 2009). In this thesis a qualitative research methodology was selected more specifically a multi-case study.

For this multi-case study, a set of cases was selected as the research object, which are the ten interviews in this study, conducted by interviewing e-commerce professionals. Multi-case studies aim to explore, describe and explain cases mainly through how and why questions (Yin, 1994). The aim is to describe systematically, accurately and truthfully the characteristics of the subject of the study (Anttila 1996; Hirsjärvi et al., 2009). It is essential that the case under study forms a kind of whole. As a multi-case study uses a variety of data collection and analysis methods, it cannot be considered as a data collection technique alone. The case study does not limit the choice of methods and both quantitative and qualitative methods can be used. Case studies are conducted in a wide range of disciplines (Yin, 1994).

In general, multi-case studies are chosen as a method when there is a need to understand the subjects in depth and to take account of the context and background. Careful examination of cases can provide insights beyond individual cases, although they cannot be used to make generalizations. The relevance and validity of the results can be strengthened by presenting a thorough description of the data and its analysis.

The empirical part was executed by performing several semi-structured interviews. In a semi-structured interview, the questions may not be always presented in the same format and they may not be precisely formulated in advance (Kallinen & Kinnunen, 2021). The literature related to the research topic was first examined, after which the key themes for the research were selected. The interview structure was then formed based on these themes. The popularity of the semi-structured interview is based on the freedom that allows the interviewee to speak freely and these kinds of interviews are also relatively easy to analyse by theme. However, it may be problematic that the themes set in advance may not be the same as those that proved to be essential after the interviews (Hirsijärvi & Hurme 2001).

The reason why interviewing was selected was because it was the most suitable way to study such an abstract research question. According to Hirsjärvi & Hurme, (2011), the interview is a flexible and suitable for many purposes. Interview is suitable as a method of obtaining information when it is not strictly defined what kind of answers will be received or when the answers are based on the interviewee’s own experience. The interviews are said also to be suitable in situations where the aim is to strive a rich description of the topic rather than just validate a selected hypothesis (Mason, 2010). Interview is considered as a particularly sensitive method as the interviewee may experience the interview as an oppressive situation and therefore leave things unsaid. On the other hand, the interviewee can also tell a modified truth which can lead to false conclusions (Hirsjärvi et al., 2009). Semi-structured interview was selected rather than structured interview because in semi-structured interviews it is easier to find new perspectives and insights because the structure of the interview is not as strict. The objective of the semi-structured interviews was to obtain insights related to the second research question that was presented in the chapter 1 of this thesis.

To reach the objective, ten semi-structured interviews were performed with professionals working with e-commerce. Basic information about the interviewees is presented in Table 4. It was up to the interviewees to choose the language in which they wanted to be interviewed. The interview form is attached in both Finnish and English and can be found at the end of the thesis as APPENDIX 1. The interviews, which were conducted in Finnish, were translated into English after transcription. Translations were made with the aim of being as verbatim as possible. These kinds of interviews are also often used when the topic is sensitive like in this case because people often don’t like to talk about possible capability shortages (Metsämuuronen, 2005). In qualitative research, the material is needed only to the extent necessary for the topic. One way to solve the question of the adequacy of the material is to use saturation. The saturation point is reached when a new interviewee doesn’t bring new information to the research and the answers begin to repeat itself (Eskola & Suoranta, 1998). In this study, ten diverse experienced interviewees from different industries were enough to create a comprehensive picture of the topic.

The interviews were analysed by using thematic analysis which can be considered as a form of qualitative content analysis and its purpose is to identify the themes that are relevant to the research problem (Eskola & Suoranta, 2008). The interviews were analysed by using qualitative content analysis that focuses on what issues, topics and themes, the material, which is analysed, is about. It can be used to analyse written texts, interviews and recorded speech but linguistic and expressive factors are excluded from the analysis. To avoid getting caught in the details, attention is paid to both differences and similarities between material units. In practice, this usually means that the material must be able to be divided into several units of analysis, which can then be compared. Before the analysis, the interview material was transcribed. According to Hirsijärvi and Hurme (2008, p. 138), the transcription of the material helps the researcher to identify common factors and themes that have emerged in the interviews. Word-to-word transcription was not performed, and unnecessary filler-words were removed. According to Hirsijärvi and Hurme (2008, p.140), this is justified, as the phenomenon under study don’t require word-for-word transcription. After that, the material was read through several times. The interviews lasted on average 38 minutes and 5 seconds. However, it should be noted that the introduction part of the questionnaire was excluded from the recording.

TABLE 4 Information of the interviewees

ID of the interviewee	Experience of e-commerce (years)	Current job title	Main industry	Turnover (in millions)
I1	12	Chief executive officer	wholesaler	10-20
I2	6	Marketing technology manager	e-commerce	10-20
I3	4	E-commerce manager	manufacture of wearing apparel	2-10
I4	10	Sales and marketing director	advertising agency	1-2
I5	4	Web analyst	management consulting	2-10
I6	4	Team lead – web analytics	advertising agency	2-10
I7	>20	Head of digital	holding company	100-200 (e-commerce's share ~10% of the turnover)
I8	4	Chief executive officer	retail	0,2-0,4
I9	4	User acquisition manager	IT consulting / IT services	N/A
I10	>5	Serial e-commerce retailer and board professional	several	several

4.2 Benefits of data analytics for e-commerce

All interviewees were initially asked to consider the benefits of analytics from an e-commerce perspective. The question is quite broad, and the benefits are not specified in more detail in the question. The perspectives of the responses varied which is also partly due to the respondents 'own domain/field related to analytics. Every interviewee felt that leveraging analytics was essential to modern e-commerce business. Two themes emerged clearly from the responses: monitoring and developing the business and eliminating the gut feeling. Many interviewees found that the benefit of analytics is knowing where the business is heading and being able to react to change. Analytics was also felt to provide answers, for example, order volume forecasting, the effectiveness of marketing campaigns, selection of sales channels and, in general, to identify cause and effect relationships.

> "It is a good monitoring tool, and with analytics we know in which direction we are going. In terms of sales and targeting, it is extremely important for the merchant to know who is buying, what customers are buying, what they are currently looking at, what kind of customers are currently liking each product, what trends, what can be sold better to the target audience. However, at best, we are a sales organization. It is highly important to know from which channel we get traffic, from which channel we have the highest conversion rate. In our streaming service we look very closely at where customers come from and how long will they stay with us." (I7)

Some respondents also highlighted the importance of everyone having an adequate understanding of the data to make things understood in the same way and to make them easier to discuss. Another important issue was considered to be that utilizing analytics eliminates guesswork and decisions made based on gut feeling. Eliminating the gut feeling and basing decisions on information provided by analytics supports decision-making that can be used to develop better business.

> "Analytics and data can give you real competitive advantage. From there, many things can be dug that can be used to develop the business. In my opinion, that's one of the most important things that should be in order. Analytics helps in the development of the business and there is no need to guess things, but on the basis of the data it is possible to say that when we decide to do something, it will work. Analytics gives such strong directions to what we are doing. People typically have their own experiences and insights into how things work, but those insights can differ a lot from what customers think." (I6)

> "If we assume that everyone working with the online store knows how data is collected and that what each data point really means, then it is a great tool because it is unbiased and able to impartially show that something seems to be so because the data shows that it is so, and all individual opinions can be stripped away. Then, of course, there may be differing opinions as to how that data should be interpreted, when one figure may be good for one and bad for another. The biggest benefit is that you get unbiased evidence on various things." (I2)

> "The main benefit is that it removes the gut feeling, which is however quite prevalent all the time. Decisions are the only ones a company has to make and if there is no data available, what will ultimately come from that raw data that you need to collect from the right sources and combine in the right ways. The second issue is enabling continuous improvement. Without data, it's hard to get into that kind of continuous thinking about how we can do things better, for example, how to serve customers or

> how to make a particular thing sell better. It is impossible to do that based on factual data if that data is not collected and available.” (I5)

At this point, the interviewee, I5, was asked to reflect on a concrete case where analytics was used to identify a problem that ultimately led to better sales figures and therefore better results. Respondent I5 works as a consultant and it did not take him long to think about the example case. The case is a model example of how a company selling services or products online can make use of analytics:

> ”Yes. For example, there was a case that was done for a company selling travel packages. They had a problem with not being able to keep people coming to the front page to look at travel packages and they would leave the page before they had had a chance to properly look at anything. That problem was first identified using web analytics [Google Analytics]. We then set out to find a solution, that this could be related to not showing enough relevant results and went on to make a change to the site itself to make it responsive to where in the world this visitor was coming from. Then that page started to dynamically change based on the information. Pretty basic stuff in the big stores, but it hadn't been done there yet. Then when it was done, it was noticed that i.e., travel packages to Tallinn started to sell much better as the page reacted to where the visitor was coming from.” (I5)

One of the respondents combined analytics into the success factors of digital commerce and described successful e-commerce and its activities as follows:

> “I'll go a little further and ask, what are the key success factors in digital commerce, so there are two things, one is your own concept, that is, a good customer promise and competitive advantage and a good brand. The other side is very good growth hacking with iterative data-driven testing at the core. And often that's what you do first to make the growth – the first million – and then the other nine million is done with a good brand and crystallizing the customer promise and everything else. Then we go back to that again, with the optimization of the data to grow more of it. That's the tip of the iceberg, but I see that as the other big area. Customer acquisition is all about growth-hacking, performance marketing and on the other hand it is a product game, i.e., constant analysis of what sells, where the sales margin comes from, where the marketing costs go, etc.” (I10)

4.3 Identified problems of data analytics in e-commerce

After being asked to reflect on the benefits of analytics, interviewees were asked to consider the weaknesses and problems associated with analytics. Again, this is a broad question, with no limits or constraints on the answers, and the interviewees were allowed to answer the question as they saw fit.

There was more variation in the answers to this question than in the question of benefits. Some respondents felt that the problem is that the data does not provide enough concrete benefits and that not enough time and resources are spent on analytics. It was also said that when there is too little time then the analytics is used too one-sided, and the numbers are interpreted too absolutely.

> “It's not monitored well enough in relation to how much time should be spent on it. They [merchants] should at least go in regularly to see what's going on. It's also not often documented. Then you also don't know how to look at the right things and they look at the wrong metrics too absolutely. One thing that customers follow narrowly is that when you look at an online store, you see the last-click transactions that came in, i.e., organic, from Google, etc. So, let’s say you get a thousand [euros] from Facebook, five thousand from Google and three thousand from paid, but what you don't see is what all contributed to that purchase decision, but you only see the last click and if you make decisions based on that alone you can mess up a lot of things. So, you follow a metric in too absolute way, and you don't see the big picture and what all the things that have contributed to the buying decision. That is missing in a large proportion of merchants.” (I4)

> “The biggest problem they [merchants] all face is the use of data. There's a lot of data and a lot of data analysis, but how to use it is by far the biggest challenge, I think. There's a lot of potential in data, but not enough time and resources are being spent on using it.” (I7)

A follow-up question was then asked as to why not enough time and resources were spent on recovery and the answer was as follows:

> “Good question, I don't think Finland is that far along yet. Finland is good at collecting data. It's a bit like marketing, that the Swedes are boldly going past us on that side.” (I7)

The responses also indicated that analytics alone is not enough. Analysis helps to understand but does not offer suggestions as to why something is not working:

> “Yes, so there are challenges, and you might not get the information you're looking for. For example, at a certain point in the purchase path, if customers fall off so if you want to know why they fell off, you can’t know that just by looking at the numbers. Analytics alone is not enough. It allows us to understand if there is something wrong, but it does not offer suggestions.” (I1)

Another problem that emerged from the responses was that there was perceived to be too much data available, which is linked to the fact that the data and analytics is ultimately not very useful.

> “The biggest problem is that there is too much data. So much that it's really challenging to analyse it consistently if you don't have the experience and skills. If you think about Google Analytics, even there you already have so much data about the digital purchase path that most retailers can't get anything out of it. So, in a way, the biggest challenge is to build the numbers that fit your business. Another key point is how to manage it and how to implement it. For a larger company, the challenge is how to get everyone involved into analytics and how to get everyone to understand analytics and how we teach people to make decisions based on analytics. Those are the two clear challenges for a small company and for a big company.” (I10)

> “If you don't know what you should be monitoring or if there is too much data, you can get so confused that you don't do much, that even if you collect it [data], you don't

> monitor it. There are certain key figures perhaps, but no one questions why this is being looked at. For example, the e-commerce bounce rate, which is a bit different from the bounce rate of email, so someone may mention that it is really good when the bounce rate of the page has fallen, but in the end, it does not say anything. What is rarely discussed is whether it makes sense to track this number and why we are tracking it. The weakness is that numbers are stared too much, and the most important things are forgotten, and people are blinded by the fact that there is too much of everything new and fancy." (I2)

Some respondents felt that the biggest problem was that there is not enough analytics expertise within their own company, which is why many have outsourced analytics. The number of skilled people in the field was perceived to be rather low and more were needed. In the same context, it was mentioned that there are more than just analytical skills and that it would be good to have other data related skills in addition to analytical skills.

> "There are quite a lot of weaknesses, but perhaps the biggest problem is that there are not enough people or not enough skilled people like my colleagues say. There are quite a few new people coming in, despite how much hype there is around data science, but you have to get those people in first. That's why quite a lot is outsourced to consultancies, for example, which is why I'm working with a lot of clients at the same time because they can't get people recruited. The other problems are related to data quality. It's quite difficult to manage and it takes a lot of things to be able to manage the data. The second is data aggregation, because nowadays that one source is not enough, you have to have it from many places at the same time or it's not trusted. Then, of course, there is the reporting of the data, so something is done with the data. And that's the biggest problem, that you can do all these things correctly, but if you don't present it in a way that you can directly say that we need to start working this way and try something new, it all just sits in databases and reports. There are some top-level problems there." (I5)

The next respondent's company had outsourced its marketing analytics to another company and was asked to explain the reasons behind the decision.

> "Outsourcing was done because we don't have that much expertise, even though I've done it myself, it's still such a world of its own and you really have to do it all the time every day to keep up with it and get really good quality results, but perhaps all the challenges related to data have been largely that we do not have the kind of data-focused expertise in-house, so all the more challenging data analytics related is still a bit of a phase. We know that we have a lot of site visitors and we have a lot of everything, and we have an enormous amount of information. We are able to use it to a certain extent at the moment, but then there is a lot that we could get more out of, but we don't know how to get it out yet. That's probably the biggest thing we often wonder about. You often hear e-commerce retailers being told that data is king, but nobody really tells you how to use it." (I3)

Two interviewees felt that the bias of advertising tools was a big problem. By bias, they meant that the ad management tools show that a particular sales event came through their ad and do not show which other events contributed to the customer's purchase decision. This bias distorts the data being analysed, making it more difficult to analyse the data and therefore, for example, more difficult to select advertising channels and budget for advertising because the data cannot be trusted. In the same context, the difficulty of combining data from different platforms was also highlighted, and for the first time the implications of GDPR and privacy were also highlighted.

> "There are plenty of them. One is that analytics is not so simple, because there are so many platforms, Google, Facebook, marketing platforms, e-commerce platforms and they don't talk to each other at the snap of a finger. Then one is attribution modelling, which is basically where a customer has come from and how they found your online store. Being able to actually track customer paths is really tricky and analytics is also misunderstood and misread. For example, if it is considered that a sale has come through Facebook, when it may be that the customer has seen a Facebook ad after the customer has already made a purchase decision. Quite a few ad management tools take credit for themselves that the sale was made through them. Because of GDPR and Apple, collecting data is and will be more difficult because people have more privacy." (I9)

> "The challenges are brought by the bias of those advertising tools, which is quite obvious. Facebook Ads shows you that all the sales come from there and Google Ads shows you that they come from there, if we exaggerate a bit. Google analytics then tries to hide that bias and there is no such unbiased narrator to tell where that transaction has come from. I say that there's a lot of opportunities, but that bias is the downside. "(I8)

The interviewees' answers were quite varied, and the responses dealt with the issue from slightly different angles, some dealing with the bias of advertising tools, some dealing with problems related to data quality and data aggregation, and some dealing with the general lack of knowledge about how to use data, even though data is collected and available in large quantities. The fact that there is so much data was seen as leading to a frequent lack of knowledge about what should and should not be monitored. This was also linked to differences in data related skills between employees in a company.

Some respondents felt that the problems of using data were due to a lack of time, but also to a lack of skills. The lack of skills was seen as a result of the fact that although the demand for data scientists has increased, there is still a lack of skilled professionals in the field, which has led to a number of online retailers outsourcing analytics rather than hiring their own data scientists.

The cost of labour and the scarcity of financial resources are certainly contributing factors. The work of a data scientist often does not bring direct financial benefits in terms of sales figures, but rather benefits in the longer term. This section highlighted the interconnectedness of analytics problems and how issues are interrelated and how many different things can cause problems. The interviewee, I6, summarized many of the problems mentioned above when answering this question as follows:

> "There are a lot of problems. Very often, the know-how on analytics is lacking and the analytics and the data is only as good as it is designed and collected. Just as it can be a

> driver, it can also steer you in the wrong direction if you don't measure the right things in the right way. In analytics, it is very important to find what is relevant and focus on that. Many times you see that clients have metrics, but they don't match the objectives, i.e. you don't measure the right things for the objectives.” (I6)

4.4 Data-driven decision making in e-commerce companies

The next theme that emerged from the interviews was the process of data-driven decision-making. Two interviewees saw the use of data in decision-making as strongly dependent on what is being decided. They also highlighted the importance of the brand and the fact that data is used more to support decisions when it comes to larger investments.

> “It's a difficult question, it depends on what is being decided and what is being discussed. There are some situations where you really dig deep: if you make big decisions that require big investments and when you go into a new business area. But for smaller things, it varies. For example, when we were going into the e-book business, a lot of data was and still is being dug up, because it is a very changing market at the moment.” (I7)

> “It depends on what you are deciding on. Things related to e-commerce, like usability and layout etc., there it will act as a strong support, but it's balancing between brand and data. I would put it so that brand first and data second.” (I3)

The next response addressed the problem already raised earlier, the inability to make decisions due to the abundance of data. With so much data available, it is not necessarily possible to make decisions based on data and it is easy to rely on the gut feeling that was already mentioned earlier.

> “There is easily too much data, making it easier to do it (decisions) by using gut feeling. Even if the data is being tracked, but the tracking is not structurally clear, the decision is still based on guesswork because you haven't been able to form a truly coherent conclusion from the data, so the decision is more like "based on all this, I think this is how it goes". Then you're on the right track, but it's still not completely data driven. The aim should be to be able to form clear KPIs.” (I10)

At this point, the interviewee, I10, also raised the importance of leadership, as the interviewee opened up his understanding of what it means to be truly data driven. The interviewee also gave a concrete example of how easy it is for a manager to make decisions traditionally without data and how decision-making processes should go in a data-driven organization compared to a traditional process.

> “A big problem in management is how we dare to give the decision making away from the leader. When someone asks me what time an email campaign should be sent out, it would be easier for me to just say, "Send it out on Sunday night at 8pm”, because it took five seconds, and it went ahead. But that's not how I create scalable organization.

> How can I trust people to make the right decisions, there's only one answer to that and that's that I can trust people's decisions in that situation when they're based on data? So, when a decision is based on data, it's no longer an opinion or a trivial decision and at that point it doesn't matter who makes the decision when the decision is based on data. In a large organization, we should move in a direction where management is no longer about making decisions, but about coaching people in decision making techniques and the use of data. Instead of telling an employee what to do, I ask what data you need to make that decision. In this example case, it would be that what were the different conversion rates in the email campaigns sent on different days. Okay good, where do you get that from? I get it by combining MailChimp and Google Analytics data. Okay, well what does the data say? Well, the data says that actually Monday morning would still be a bit better than Sunday evening. This is what should be used to create scalable organization." (I10)

Respondent I2 felt that the overall level of data driven decision making in their company was good compared to many other companies. However, the respondent felt that the level of understanding in the company should be better, as currently the wide differences in skills between the different parties involved lead to problems, as it is difficult to discuss data-related issues if not all parties have a sufficient understanding of the topic.

> "The biggest problem we have is that you can't expect everyone to understand data, but I bet we're still really good at making decisions based on data. There are always weaknesses, but I've seen companies where data skills are almost zero. For us, it's the lack of knowledge, because not everyone speaks the same language, so it requires from those who understand to sometimes spoon-feed what each figure means. It requires mutual understanding. You can't assume that everyone is deeply involved in every data point, which can easily create a situation where someone is a bit of a 'data snob', which creates a situation where you don't want to talk about it when you are at such a different level, even though you should be talking about it. It has to do with understanding data as well as understanding people." (I2)

The interviewee was then asked to describe an example situation where analytics was involved. The situation described illustrates well how the information from the data can also be used in a situation where a change was made, but the data showed that the change was bad, so the original situation was reverted to. The example situation described by the respondent was as follows:

> "One example was that we did a rebranding in autumn 2020 and changed our front-page communication quite a bit. By front page I mean the page that you land on if you type the name of our store into a search engine. Then we noticed that the traffic that was coming through the front page, their conversion rate had dropped significantly when compared to a month earlier. Of course, there could be many other reasons for that, but one clear variable that had happened was this change in messaging on the front page. We made the change back and made the page more aggressive and that brought the conversion back up to the same level as it had been. That's one concrete good example." (I2)

One interviewee felt that the customers often lack the courage to try new things and the confidence to trust the data. On the other hand, it was also pointed out that some try to measure everything and stare at the data in too much detail.

> “I think the biggest challenge is not to dare to make decisions based on analytics, to use too much gut feeling. We've had cases where a client is getting a ridiculous return from a particular advertising channel, and it would be worthwhile to start increasing the amount of advertising space gradually and see how long the return remains good enough in relation to the advertising investment. In other words, if you get a tenner for every euro you spend, shouldn't you increase it as long as the ratio remains acceptable. Let's say that there are almost few customers who dare to do that, even though there is basically no risk, because you can stop buying that media space at any time. People do not have the courage to make enough use of that information. Another counterweight is those who measure too much and think that everything must be measurable. It's easy for me to say, because I used to do digital marketing and I used to think that you always have to be able to measure everything, but when you have a lot of marketing and sales channels that are difficult to measure, but they can still be a good thing. That kind of over-measurement can also be bad, but it's less common in Finland” (I4)

Overall, respondents felt that the use of data in decision-making was a very important part of the process. The responses indicate that data is seen as a tool to help justify decisions and assess the consequences of decisions. The use of data in decision-making was seen as strongly dependent on the issue to be decided and on the size of the investment, if the decision involves financial investment. Again, the responses highlighted how much easier it is to make decisions based on a gut feeling and previous experience. Too much data was also felt to influence decision-making in a negative way. Lack of courage, lack of trust in data and large differences in skills within the company were also highlighted as barriers to data-driven decision making. One interviewee also raised his views on what data-driven decision making requires from a leader and offered a practical example of that.

4.5 Capabilities of case companies and their evaluation

This chapter reviews the interviewees' responses to the capabilities. The interviewees were asked what kind of resources their company needs to use data analytics and in which areas they perceive gaps. The answers are structured in the same order as in the literature review. Firstly, we review the answers related to tangible factors, followed by human factors, and finally we look at the answers related to intangible factors. As can be seen from the responses, an absolute classification of the responses is impossible because of the interdependence of the different factors. However, the responses clearly show the factors where respondents perceived shortcomings as well as the factors that respondents perceived to be in good shape.

4.5.1 Tangible factors identified through interviews

There were surprisingly few responses on basic resources. Money was not perceived as a limiting factor in any other context, except that some respondents felt the need to hire more analytics experts, but they did not feel that this made sense at the moment due to the small size of the company.

> "Our team is really creative and kind of a big-picture type of team, that we all like to develop new things, but the details are not our strength, that I could see anybody from the current person becoming a data scientist person, but I see it as particularly important that we have such a person in the future, but maybe not yet. It's at a time when it's better to take advantage of the skills we have now and put in place tools that I could use, given that I have some understanding of the issues myself. Then when we get to the next growth targets, then we could think about having a separate person who would just focus on analytics. I follow a lot of other online shops, so I know that they have separate analytics teams, where they track and manage things, so I can see that it's an absolutely important part of the future." (I3)

Time was seen as a more limiting factor than investment among respondents. In addition, I2 attributed the lack of time to the fact that it limits the training of employees and thus their ability to increase their own technical skills. Time was therefore also linked to organisational learning and technical skills, which will be discussed further in later subsections.

> "Yes, it would be time. For example, if you're doing tests on a page, you can't have many different tests running at the same time, otherwise it will give you an incorrect picture of what worked and what didn't. So, there can't be too much overlap, which contributes to some delay, but very little because we have quite a lot of traffic on our page and we can get comparable results in a couple of weeks and be able to run another test." (I1)

Understandably, there were more data-related responses. Based on the literature review, data-related problems could have been related to data accessibility, accuracy, quality, quantity, velocity, variety and merging data, but respondents felt that data availability in particular was very good. Many respondents felt that the problem was that the data was not always collected correctly, leading to misinterpretations. This was also felt to affect the quality and reliability of the data. The fact that the data is in many different places was also perceived as a problem, as many people do not know how to combine the data in the same place in a way that makes it usable.

> "I would say that if you look at a single data point, availability is pretty good these days. I would argue that availability is not the big problem, but whether the data is collected correctly in the first place. It's really easy nowadays to collect the data yourself or get it from other tools, but the problem is whether it's collected correctly. For example, there are so many things you can do wrong on a website and often they are discovered too late." (I5)

> “There is often too much data, but the quality of the data is often the problem, whether we have been able to measure the right things, whether it is reliable, whether it has integrity and whether there are gaps. That's usually where the problem lies. There is usually data, but the quality is the decisive factor in the end, and it is usually incomplete in one way or another.” (I6)

> “There are also challenges in that the data is in many different places and although there are solutions to bring it all together, it is not being done. In e-commerce, the channels include google analytics, the e-commerce store's own customer data, then there may be hot chars or similar, and you should also monitor the data from advertising tools such as Facebook and Instagram. Maybe also brand and page data to see how the pages have evolved. Then there's email marketing and automation software with data. Then there's Google Ads probably and there's data there as well. So, there's a lot of data and you should understand what's important and what's worth monitoring. There's some software that's been done for clients that aggregates the data for the client into one view. I can also see that much more could be used in the decision-making process and that people should dare to make decisions and try things out, as long as the changes are documented. In other words, there should be more of this kind of growth hacking in companies.” (I4)

The response from interviewee I4 therefore also highlighted the link to decision-making and how decision-makers should dare to experiment more. Respondent I10 disagreed with the previous respondents and generally considered all tangible factors to be manageable.

> “Investments, time, data, data availability, etc. are all manageable. It may be something that many companies say is a problem, but it's a bit of an easy way out to say that it's a problem when it's not. You can get Google Data Studio for free and if you want to integrate it with ERP, all it takes is watching Youtube videos and learning. Any half intelligent person can get it done and is it is worth the investment of time – it is worth the investment of time. So, tangible factors are not directly a problem but an excuse.” (I10)

Based on the results of the literature review, problems with information systems can be related to, for example accuracy, reliability, security or confidentiality. However, little else was associated with this context other than the e-commerce platforms often used by online retailers. Only a few of the most common platforms were mentioned, such as Shopify, Magento, WooCommerce and BigCommerce. The platforms were not seen as a limiting factor and the interviewees mainly mentioned only individual shortcomings and features related to specific platforms.

> “Shopify's biggest challenge is the customization of the checkout path, that it's a bit rigid, but we didn't want to go down the customization route because that would then prevent the benefits of SaaS, but I think that's the biggest weakness.” (I7)

> “In modern platforms, there are certainly not that many limiting factors, if it's hard coded from scratch yourself, then there can be challenges with measurement, for example. Often for the very reason that it has not been taken into account from the

start. Shopify, WooCommerce and Magento, there are very rarely problems with them." (I6)

"We are a bit biased because we make e-commerce solutions on WooCommerce platform. But for almost all platforms that are even a little bit up to date, you get that Google Analytics level of tracking. […] I think good ones are WooCommerce, Shopify and if you have more money, then Magento. For BigCommerce, there are not many experts in Finland. I would almost leave out the others for SMEs." (I4)

In relation to the e-commerce platform, interviewee I9 pointed out that no platform is ready immediately and that an e-commerce platform alone is not enough. In the same context, I9 referred to the importance of technical skills, which we will discuss in more detail in the next subsection.

"A platform is important, but no platform is complete, and no e-commerce platform alone is enough. It would be like building a house with only walls, you can't live in it yet. Yes, you need a lot around it, and it requires a lot of technical know-how or a lot of capital to outsource the technical stuff." (I9)

Table 5 below shows the summary of the results of the interviews related to tangible resources. The first column indicates the type of capability into which it is classified, in the same way as in the literature review results. The second column shows the resource identified from the literature review which was discussed. The third column contains descriptions of how each factor was perceived by respondents. Same columns are used also in Table 6 and Table 7.

TABLE 5 Summary of interviews classified according to the results of the literature review – tangible resources

TYPE OF THE CAPABILITY	**Identified resources based on literature review**	**Key findings of how different factors were perceived by respondents**
TANGIBLE	Basic resources (investments & time)	• Investments was not perceived as a limiting resource except that some respondents felt the need to hire more analytics experts, but they didn't feel that this made sense at the moment due to the small size of the company. • Time was seen as a limiting resource among respondents. In addition, according to I2, the lack of time limits the training of employees and thus their ability to increase their technical skills and knowledge.
	Data	• Respondents felt that the problem was that the data is not always collected correctly, leading to

		misinterpretations. This was also felt to affect the quality and reliability of the data. • Data siloing was also perceived as a problem, as many people do not know how to combine the data in the same place in a way that makes it usable.
	Information systems	• The e-commerce platforms were not seen as a limiting resource and the interviewees mainly mentioned only individual shortcomings and features related to specific platforms.

4.5.2 Human factors identified through interviews

This subsection presents the responses related to human factors. The responses are categorized according to the findings of the literature review, although some of the responses are also linked to other areas.

Business knowledge was only explicitly referred to in the responses of two interviewees. Interviewee, I2, mentioned that they are currently recruiting someone with an understanding of the customer and data. He stressed that no matter how sophisticated the data is processed, it does nothing if it doesn't lead to action. This can be seen as a link to the decision-making process, which was already discussed in the previous chapter.

> "We talked in recruitment about looking for people who have both that business understanding and also a little bit of data understanding, but that the business understanding is more important, because it's easier to teach someone the basics of analytics or a little bit deeper." (I2)

Interviewee I10 gave a comprehensive answer on the same topic, discussing the reasons why outsourcing is not really a solution. In addition, he compared the Swedish businesses to Finnish businesses and described the reasons why there are more really successful international B2C e-commerce businesses in Sweden.

> " It's really the people that are the problem, it's really about finding the right people. Buying from outside is not a solution either, because then you're buying from a data consultancy, but that's not the best solution for e-commerce company. [...] In e-commerce, you should do direct tactical data analysis to make more sales next week. It requires a combination of commercial thinking and data understanding and continuous testing, which is related to the actual data benefit, because only through growth hacking you get the benefit out of it. You should be able to make different assumptions based on the data and go out and test them and do continuous iterative testing. It requires that you are inside the business. [...] So in other words, that expertise should be brought in-house. The problem is that this kind of data-driven expertise

with a commercial view is not really there. [...] In Sweden, for example, there are great growth stories where each marketing channel is actually monitored with different attribution models and the benefits are dug out from there, to see how much marketing effort is being invested in this channel and that channel. That kind of has not really been done in Finland and no one knows how to do it, because it is not trained anywhere except in those companies where it is done well. The reason why this has become part of the culture in Sweden is that there are many companies that have already done it well and if you want to set up a similar case with a large-scale investment, you go to see who previously did the analytics side and hire the person who has the skills. That's how you get that culture transferred from one firm to another. There is not much of this in Finland yet." (I10)

All respondents referred at some point in the interview to a lack of technical understanding and skills and saw this as a clear problem. The responses also highlighted the large differences in skills between the company's employees. Interviewees were also asked to reflect on ways to better embed analytics and related understanding throughout the company. Interviewee I6 also stressed the importance of analytics being a genuine part of the company's strategy.

"I see that the number one capability comes from that competence, because then it's much easier to be aware of what these things are needed, what kind of software should be used and what you should go even further to build - your own skills or the skills of the organisation. And it's not enough for one person to have the right skills, many people need to have them." (I2)

"The biggest is that you are measuring the wrong things. Usually, you want to do something or improve something, but you can't do that when you're measuring the wrong things. Then there is also the lack of skills in analysis, in that even though there is a lot of data, there is no knowledge of how to segment or analyse it in order to make sense of it. These are the biggest factors that do not translate into concrete action. Analytics should be part of the strategy and there should be a person who is focused on it and on top of the situation. In general, the digital world is so changing that you have to keep up with it. In general, companies lack the resources, time and skills to focus on analytics." (I6)

Interviewee I5 offered a concrete way to develop analytical skills by describing his own learning process as follows:

"I did it on three levels. I started with Google's free courses and other free material, which is really abundant online in English. The second is that you have to do things related to analytics and you have to do quite a lot of them. You can only get to a certain level with videos and reading. The third is that you have to network with people who have already done more of it than you and then you start to scale up your learning. I got the biggest benefit from getting into my current job with people who have been doing this for over a decade and that's where I benefitted the most. I would say that there are no other levels now. All people should first understand some and then do and then bring in more knowledgeable people even if it's from outside. I haven't done things in-house myself, so I may not be impartial, but in-house is a very different world because the consultant has to know a different level of these things." (I5)

Respondents also highlighted the importance of management skills and understanding, and their experiences of management attitudes towards analytics. In this context, there were very similar responses to those in the previous section on data-driven decision making.

Interviewee I4 felt that there are quite often gaps in management knowledge and understanding, and also raised the importance of documentation. I4 serves other companies and has certainly, as an outsider, seen the way management operates differently from other respondents who based their views on the situation in their company.

> “Many lack the understanding to look at and analyse and use data and make decisions. So, they don't have the skills to use the data, i.e., they don't have the know-how to make use of it. Also, as I have already said, looking at things in black and white, not documenting and thinking about other options that could have an impact. Data is also sometimes overused. I would also stress documentation.” (I4)

> “Fortunately, we have the fact that if you can prove it through data, so it is trusted that it's more about how you communicate it. There is no scepticism. Rather, if a person has a certain issue is a matter of the heart and we are sure of it, then we do not always listen so much to the data, but that does not happen too much. It's not always a bad thing that sometimes someone is stubborn and pushes their own view through. We're on the right track.” (I2)

> “I have not yet been in a position where I have to justify an investment of tens of thousands of euros, but I also don't see that it is a problem at the point where it is necessary. […] In our company, everybody understands the importance and the benefits.” (I3)

I1 is part of the management team in their company and felt that the whole management team was comfortable with analytics because it helps to understand causes and consequences.

> “No, I've always been analytical. Causes and consequences. This suits my nature, and our management likes to have a lot of information. The more information the better.” (I1)

TABLE 6 Summary of interviews classified according to the results of the literature review – human resources

TYPE OF THE CAPABILITY	**Identified resources based on literature review**	**Key findings of how different factors were perceived by respondents**
HUMAN	Business knowledge and business analytics	• Respondents perceived that it is hard to find talented people with data-driven expertise combined with a commercial understanding.

	Managerial skills	• Some respondents felt that there are quite often gaps in management's knowledge and understanding in analytics, but some disagreed on this point and felt that analytics brings certainty and support to decisions, at least in their company.
	Technical knowledge and technical skills	• All respondents referred at some point in the interview to a lack of technical understanding and skills and saw this as a clear problem. The responses also highlighted the large differences in skills between the company's employees.

4.5.3 Intangible factors identified through interviews

This subsection presents the responses related to intangible factors. The responses are categorized according to the findings of the literature review and again some of the responses are also linked to other areas. There were fewer direct responses to intangible factors than to other types of factors. Governance was the only factor identified in the literature review results for which there were no direct responses during the interviews. Only one respondent gave a direct answer on the data-driven culture, and this respondent gave a more extensive answer in the interview. Interviewee I10 felt that the prevailing culture was the reason why Finnish e-commerce companies often focus only on Finnish market. He cited finance as a reason for this. According to I10, Finnish online shops are less likely to have investor financing, but instead build their businesses with their own capital and a corporate loan from a bank, i.e., Finnish online shops rarely have access to so-called risk capital. However, he saw that this is changing.

> "It's the culture and there's too much focus on Finland. One of the reasons why there is too much focus on Finland is that there hasn't been a culture of funding on the field of e-commerce and also the level of ambition hasn't been so high. You can't say that it's badly done or that there are bad e-commerce companies, but if you get the financing side involved, which comes from the Slush and start-up scene, you immediately start building a billion-dollar company, and when you start doing it through financing, you have to build a billion-dollar company, because the venture capital investors wouldn't even have got involved if the goal was so modest. In a way that's partly an excuse, but in a way it's the real reason. When you go 100% with entrepreneurs' own money, the level of ambition is not so high. In a generally internationalising digital world, every single business is international, so the primary idea should be how to make the world's best online store, not how to make Finland's best online store. I also believe that that culture is slowly changing." (I10)

After the previous answer, interviewee I10 was asked what should be done differently to make Finnish e-commerce companies more international and the answer was as follows:

> "The level of ambition, even if it's not related to knowledge management, but it's related to the general business culture, to get the mindset that gaming industry companies have, that when a couple of guys graduate from games studies and start a company, none of them are going to make the best mobile game in Finland - that's nobody's goal. Everyone's goal is to make the next blockbuster game or the best thing in some niche scene. Of course, you have to remember that when ambition is high, the risks are higher, some cases fail, and few succeed [...]" (I10)

Decision making and the problems associated with it were already discussed in more detail in section 5.3, but here are some more thoughts from interviewees on the main reasons why analytics often fall short of measurement and still fail to make decisions based on the results of analysis. I5 cited the perception that analytics is very technical and the rush to produce final reports due to the high workload, which results in low action and therefore low impact. Interviewee I4 identified the lack of systematicity, knowledge, understanding, regularity and courage of decision-makers as a reason.

> "I would say that one aspect is that it's seen as really technical and there's a lot of focus on the technical side, but not on how to present all the effort in the end. If it falls apart at the end and the listener is not interested, then all that groundwork has been for nothing. One reason is that it is left to the fact that all this is of high quality, but that the final report is done in a hurry. That may be one reason, at least from what I've seen outside my company, for how in-house analytics might work. It's probably because there's a lot going on at the same time and those real actions are being taken are small." (I5)

> "Small things, without exception, would go a long way in making things more regular and systematic. Try to set more targets and do more experiments based on analytics and thereby improve performance. In other words, more courage to make decisions based on analytics. Then you don't even necessarily understand the possibilities. It seems that if you have your own money tied up in it, it motivates you more, but many people lack the motivation to develop their business. We have a lot of customers who have a core business and e-commerce has become a new thing, but there is a lack of knowledge and understanding of e-commerce and at the same time managers are afraid to invest in it." (I4)

Responses on organizational learning and innovative culture related to training opportunities in companies, attitudes towards new innovations and practices. Interviewee I2 also mentioned that more time for training would be needed in order for people to learn new analytics-related issues during working hours. I2 also mentioned that this always costs the company a certain amount of money. I5 works as a consultant and felt that attitudes correlate significantly with how much real benefit the company buying the service actually gets from the service.

“The difference in attitudes is staggering, that even for our customers who pay for the service, their attitudes can be night and day. Some buy because it's a must for a big company and some buy because they feel it's a very important part of our overall business, but still don't want to create their own team. Some customers may have an internal culture that data is seen as a threat that will take jobs, while another company may have a culture that they can't remember a time when analytics would not have benefited their work. So, there's a huge difference. I haven't seen a study that's drawn together, but I would argue that those who value it get more out of it the more they value it. Those who buy because they have to, then the benefit may even be negative, because they pay to have the processes done, but they don't benefit from it at all. To put it bluntly, this is how it goes.” (I5)

“We would need a more precise strategy for the use of data, which would then filter through all the other skills that are needed, so that the organization can learn through courses and get more time for it. And to encourage people to take the time to analyse and learn certain things. Then, that the management also understands that money has to be invested in this, that this costs money and is not free.” (I2)

Interviewee I7 works in the largest company of the interviewees. According to I7, they offer training opportunities, but not systematically.

” Sometimes we have trainings, but nothing organised. We have increased awareness and culture in areas that are not interested in analytics. The traditional side we are doing well within the group. Traditional retail data goes there, but this digital data is alien to many and that's the awareness we are trying to raise.” (I7)

TABLE 7 Summary of interviews classified according to the results of the literature review – intangible resources

TYPE OF THE CAPABILITY	Identified resources based on literature review	Key findings of how different factors were perceived by respondents
INTANGIBLE	Data-driven culture	• Respondent I10 felt that the prevailing culture was the reason why Finnish e-commerce companies often focus only on Finnish market. He cited finance as a reason for this. According to I10, Finnish online shops are less likely to have investor financing, but instead build their businesses with their own capital and a corporate loan from a bank, i.e., Finnish online shops rarely have access to so-called risk capital.
	Decision-making	• Respondents felt that the use of data in decision-making is a very important part of the process. The responses indicate that data is seen as a tool to help justify decisions and assess the consequences of decisions.

		• The use of data in decision-making was seen as strongly dependent on the issue to be decided and on the size of the investment, if the decision involves financial investment. • Respondents highlighted how much easier it is to make decisions based on a gut feeling and previous experience. • Too much data was also felt to influence decision-making in a negative way. • Lack of courage, lack of trust in data and large differences in skills within the company were also highlighted as barriers to data-driven decision making.
	Governance	N/A
	Organizational learning and innovative culture	• Responses related to training opportunities in companies, attitudes towards new innovations and practices. • Interviewee I2 also mentioned that more time for training would be needed. • I5 felt that attitudes correlate significantly with how much real benefit the company buying an analytics related service actually gets from the service.

5 DISCUSSION AND CONCLUSIONS

This chapter presents the conclusions of the study and answers the research questions. Finally, the success and limitations of the study are assessed and possible areas for further research are discussed.

As outlined at the beginning of the study, the interest towards data analytics has been rising exponentially, but the capabilities that companies need to benefit from data analytics have not been considered enough in the context of e-commerce. To fill this gap, the following research questions were selected, which aimed to provide answers to the research problem. The research questions were as follows:

- What capabilities are required for companies to benefit from data analytics from the perspective of resource-based view?
- How are the different capabilities related to data analytics perceived in the case companies?

The first research question was answered through a literature review and the second one was answered on the basis of empirical results.

The results of the literature review show that analytics and its utilization in an organization involve several different factors, some of which are more abstract than others. It is clear that each e-commerce company is unique, and each e-commerce company has its own strengths and weaknesses, which is why the empirical part is an integral part of this thesis. However, literature review's results managed to identify broadly all the key factors involved in leveraging the company's analytics. In summary, the literature review summarized that a company needs capabilities in every subarea in order to be able to leverage analytics in its own operations. In the empirical part, the aim was to find out, with the help of a literature review, which of these identified capabilities are emphasized and in which areas deficiencies in Finnish e-commerce companies are identified.

5.1 Conclusions

The study aimed to determine the current state of analytics use in Finnish e-commerce businesses and to identify factors that limit the use of analytics in businesses. The literature review identified ten different resources that make up the data analytics capabilities of a company. These ten factors were categorized into three different parts using the classification of Gupta & George, (2016): tangible, human and intangible. In addition, ten semi-structured interviews with industry professionals were conducted to describe the current situation of Finnish e-commerce businesses.

Three tangible resources were identified through the literature review that were basic resources, data and information systems. In this case, the basic resources of the firm consist of time and financial investment, i.e., the firm's ability to invest. Somewhat surprisingly investments were not perceived as a limiting resource except that some respondents felt the need to hire more analytics experts, but they didn’t feel that this made sense at the moment due to the small size of the company. Unlike investment, time was seen as a limiting resource among respondents. In addition, according to I2, the lack of time limits the training of employees and thus their ability to increase their technical skills and knowledge. According to the results of the literature review, data as a resource is influenced by access to data, accuracy of the data, the amount of data, quality of data, quantity of data, velocity and variety of data and how the company is able to merge internal and external data. Based on the results of the interviews data siloing was also perceived as a problem, as many companies do not know how to combine the data in the same place in a way that makes it usable. Respondents also felt that the problem was that the data is not always collected correctly, leading to misinterpretations. This was also felt to affect the quality and reliability of the data. The results of the literature review can be said to be at least partly in line with the responses of the interviewees on data-related issues. However, data was not considered to be a very problematic resource due to its easy accessibility and convenient analysis tools. For information systems, the results of the literature review emphasised the importance of accuracy, timeliness, reliability, security and confidentiality of the systems. Interviewees didn’t really mention problems with the systems and also the e-commerce platforms were not seen as a limiting resource and the interviewees mainly mentioned only individual short-comings and features related to specific platforms.

Three human resources were identified in the literature review, which are business knowledge and business analytics, managerial skills and technical knowledge and technical skills. The literature review identified the knowledge of the field as an important factor and also that companies need to have clear understanding of the data and analytics to benefit from them. In the interviews, respondents perceived that it is hard to find talented people with data-driven expertise combined with a commercial understanding and considered understanding of the industry to be particularly important.

In terms of managerial skills, the literature review identified the following issues as important. Management needs to have capability to optimize decision-making models. Managers need it skills to gain sustainable competitive advantage. Manager's need ability to request and consume data analyses. Manager's need business acumen and understanding of the data to know how to use it. Leadership was also seen as an enabler of change. Respondents felt that there are quite often gaps in management's knowledge and understanding in analytics, but some disagreed on this point and felt that analytics brings certainty and support to decisions, at least in their company. It can be said that managers skills in using and understanding data varies widely between companies, but both the literature review and the interviews suggest that management skills are important for the use of analytics. According to the results of the literature review, the technical skills required include skilled data scientists that have good coding skills and can use sophisticated methods and strategists who understand how to deploy the tool and supporting technology personnel to implement the data and tacticians to organize and manipulate data into operational models. All respondents referred at some point in the interview to a lack of technical understanding and skills and saw this as a clear problem. The responses also highlighted the large differences in skills between the company's employees. Technical skills can therefore be considered a very important resource in the use of analytics, but based on the interviewees' responses, it is easy to learn analytics skills on your own and many respondents did not consider outsourcing as a good solution, at least for a small online store.

Four intangible resources were identified from the literature review, which are slightly more abstract than the previous resources. The first one of these four resources is data-driven culture. Based on the results of the literature review, it was associated with analytics driven management culture and the desire to seek how to benefit from data analytics. Among the interviewees, only one respondent directly perceived this as a problem, and he also perceived the culture of financing as the biggest problem in e-commerce. Respondent I10 felt that the prevailing culture was the reason why Finnish e-commerce companies often focus only on Finnish market. According to I10, Finnish e-commerce companies are less likely to have investor financing, but instead build their businesses with their own capital and a corporate loan from a bank, i.e., Finnish e-commerce companies rarely have access to so-called risk capital.

The second one of these four resources is decision-making. Based on the results of the literature review decisions should be made based on data rather than institution and companies should have predefined strategy for using data analytics. Interviewees were in agreement with the findings of the literature review. Respondents felt that the use of data in decision-making is a very important part of the process. The responses indicate that data is seen as a tool to help justify decisions and assess the consequences of decisions. Also, the use of data in decision-making was seen as strongly dependent on the issue to be decided and on the size of the investment, if the decision involves financial investment. Respondents also highlighted how much easier it is to make decisions based on a gut feeling and previous experience. Too much data was also felt to influence decision-making in a negative way. Lack of courage, lack of trust in data and large differences in skills within the company were also mentioned as barriers to data-driven decision making. The literature review also highlighted the importance of governance in managing data and ensuring trust, but this was not mentioned in any of the interviews. As a final resource, the literature review identified organizational learning and innovative culture to which was attached the ability to share knowledge, ability to train required skills when needed, data analytics as a part of everyone's job as an ongoing process and learning about data analytics across the organization. Interviewees were largely in agreement with the findings of the literature review and one mentioned that more time for training would be needed. Respondents also felt that attitudes play an important role in how well analytics and, for example, purchased analytics services are benefitted.

Outside the results of the literature review, the interviews revealed only two other influencing factors: the size of the company and the age of the company. In fact, company size and age do not directly affect analytics or the use of analytics, but larger and older companies are more likely to have more data accumulated and generally have more financial resources to make larger investments in analytics.

The results show that the ten elements of the literature review are well aligned with the interviewees' responses, except for governance, which did not come up in the interviews. The main problems identified by interviewees was the low level of technical skills and the difficulty in finding employees with a good understanding of the business and a good understanding of analytics and how to develop them. In addition, the interviews revealed a lack of data-driven decision making. One identified problem was that the data is not always collected correctly which leads to misinterpretations, but this also relates to the lack of technical skills. The importance of everyone having some understanding of data was also highlighted. Material factors such as money, time, data and e-commerce platforms were mostly seen as neutral factors and were not seen as a problem in the development of analytics. The Figure 4 below illustrates the results in graphical form.

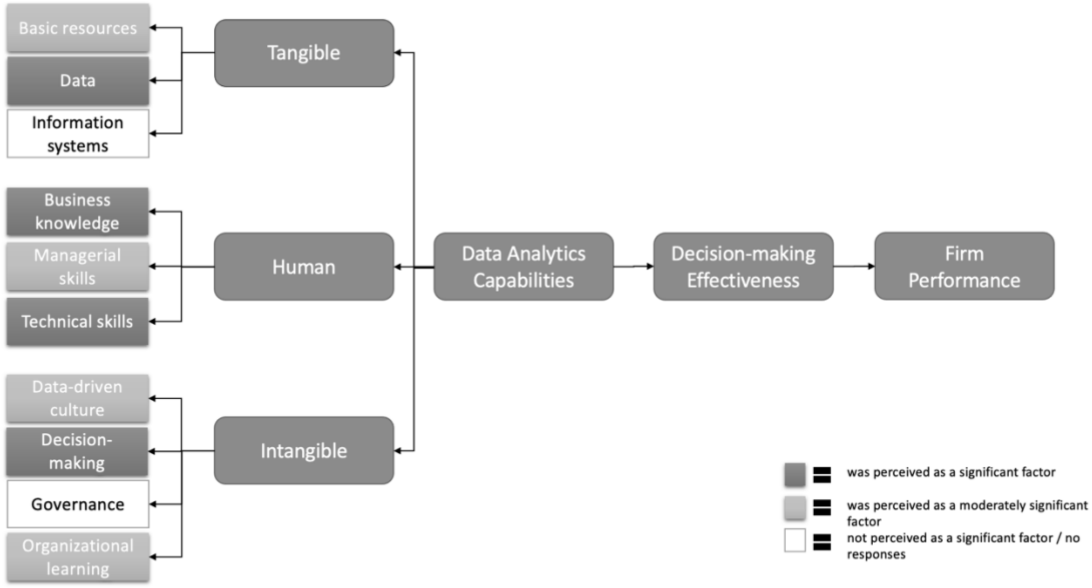

Figure 4 Conceptual framework

5.2 Limitations and future research

This study provides a comprehensive picture of the current state of Finnish e-commerce companies, and the interviewees were mainly along the same lines and identified the same phenomena as the results of the literature review, with a few exceptions already mentioned. However, it is important to understand that the study was limited to the situation of only ten Finnish companies and therefore cannot be generalised without further research. Further research could also take into account issues not covered in this study, such as issues related to artificial intelligence, implementation of analytics tools or data protection.

REFERENCES

Acito, F., & Khatri, V. (2014). Business analytics: Why now and what next?, 57, 565-570

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systemat-ic review and agenda for future research. Electronic Markets, 26(2), 173-194.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment?. International Journal of Production Economics, 182, 113-131.

Anttila, P. (1996). Tutkimisen taito ja tiedon hankinta. Helsinki: Akatiimi.

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of management, 17(1), 99-120.

Barney, J. B., Ketchen Jr, D. J., & Wright, M. (2011). The future of resource-based theory: revitalization or decline?. Journal of management, 37(5), 1299-1315.

Barney, J., Wright, M., & Ketchen Jr, D. J. (2001). The resource-based view of the firm: Ten years after 1991. Journal of management, 27(6), 625-641.

Barrett, M., Davidson, E., Prabhu, J., & Vargo, S. L. (2015). Service innovation in the digital age. MIS quarterly, 39(1), 135-154.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. Harvard business review, 90(10), 78-83.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. Review of general psychology, 1(3), 311-320.

Behl, A., Dutta, P., Lessmann, S., Dwivedi, Y. K., & Kar, S. (2019). A conceptual framework for the adoption of big data analytics by e-commerce startups: a case-based approach. Information Systems and e-Business Management, 17(2), 285-318.

Brinkhues, R., Maçada, A. C., & Casalinho, G. (2014). Information management capabilities: Antecedents and consequences.

Bose, R. (2009). Advanced analytics: opportunities and challenges. Industrial Management & Data Systems.

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

BuiltWith. (18.2.2021) Ecommerece usage distribution in the top 1 million sites. Retrieved 2.3.2021 from: https://trends.builtwith.com/shop

Carmichael, F., Palacios-Marques, D., & Gil-Pechuan, I. (2011). How to create information management capabilities through web 2.0. The Service Industries Journal, 31(10), 1613-1625.

Chae, B. K., Yang, C., Olson, D., & Sheu, C. (2014). The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective. Decision support systems, 59, 119-126.

Chandrasekaran, S., Levin, R., Patel, H., & Roberts, R. (2013). Winning with IT in consumer packaged goods: Seven trends transforming the role of the CIO. McKinsey & Company, 1-8.

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

Chen, S. C., & Dhillon, G. S. (2003). Interpreting dimensions of consumer trust in e-commerce. Information technology and management, 4(2), 303-318.

Davenport, T. H. (2006). Competing on analytics. Harvard business review, 84(1), 98.

Davenport, T. H., Barth, P., & Bean, R. (2012). How'big data'is different.

Davenport, T. H., Harris, J. G., Jones, G. L., Lemon, K. N., Norton, D., & McCallister, M. B. (2007). The dark side of customer analytics. Harvard business review, 85(5), 37.

Davenport, T. H., Harris, J. G., & Morison, R. (2010). Analytics at work: Smarter decisions, better results. Harvard Business Press.

Davenport, T., & Harris, J. (2017). Competing on analytics: Updated, with a new introduction: The new science of winning. Harvard Business Press.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340.

Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. Decision Support Systems, 55(1), 412-421.

Devaraj, S., Fan, M., & Kohli, R. (2002). Antecedents of B2C channel satisfaction and preference: validating e-commerce metrics. Information systems research, 13(3), 316-333.

Drucker, P. F. (1967). Effective Decisions. Harvard University. Graduate school of business administration. Page 2.

Eskola, J., & Suoranta, J. (1998). Johdatus laadulliseen tutkimukseen. Tampere: Vastapaino.

Eskola, J., & Suoranta, J. (2008) Johdatus laadulliseen tutkimukseen. Tampere: Vastapaino.

Freeman, M. (1984). History, narrative, and life-span developmental knowledge. Human development, 27(1), 1-19.

Frost, R., Strauss, J., (2013). E-marketing. Pearson Prentice Hall. Upper

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International journal of information management, 35(2), 137-144.

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision-making performance. The Journal of Strategic Information Systems, 27(1), 101-113.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049-1064.

Heisig, P., Suraj, O. A., Kianto, A., Kemboi, C., Arrau, G. P., & Easa, N. F. (2016). Knowledge management and business performance: global experts' views on future research needs. Journal of Knowledge Management.

Hill, S., Provost, F., & Volinsky, C. (2006). Network-based marketing: Identifying likely adopters via consumer networks. Statistical Science, 21(2), 256-276.

Hirsijärvi, S. & Hurme, H. (2008). Tutkimushaastattelu: Teemahaastattelun teoria ja käytäntö. Gaudeamus Helsinki University Press.

Hirsjärvi, S. & Hurme, H. (2001) Tutkimushaastattelu: teemahaastattelun teoria ja käytäntö. Helsinki: Yliopistopaino.

Hyvönen, K., & Pylvänäinen, E. (1999). Kirjaostoksilla verkossa. Kuluttajatutkimuskeskus.

Intezari, A., & Gressel, S. (2017). Information and reformation in KM systems: big data and strategic decision-making. Journal of Knowledge Management.

Kallinen T. & Kinnunen, T. (2021). Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. Retrieved 6.4.2021 from: https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuksen-aineistot/haastattelut/

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., ... & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. The scientific world journal, 2014.

Kiron, D., Prentice, P. K., & Ferguson, R. B. (2012). Innovating with analytics. MIT Sloan Management Review, 54(1), 47.

Koirala, P. (2012). What is big data analytics and its application in E-commerce.

Kozlenkova, I. V., Samaha, S. A., & Palmatier, R. W. (2014). Resource-based theory in marketing. Journal of the Academy of Marketing Science, 42(1), 1-21.

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. International journal of information management, 34(3), 387-394.

Lamba, H. S., & Dubey, S. K. (2015, September). Analysis of requirements for big data adoption to maximize IT business value. In 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions) (pp. 1-6). IEEE.

Lampikoski, K., & Lampikoski, T. (2000). Kuluttajavisiot: näköaloja kuluttajakäyttäytymisen tulevaisuuteen. WSOY.

Langley, A., Mintzberg, H., Pitcher, P., Posada, E., & Saint-Macary, J. (1995). Opening up decision-making: The view from the black stool. organization Science, 6(3), 260-279.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. MIT sloan management review, 52(2), 21-32.

Lawless, M. (2014). Predictive analytics: An opportunity for better demand planning and forecasting. The Journal of Business Forecasting, 33(4), 44.

Lucker, J., & Guszcza J. (6.6.2012). 5 Reasons Analytics Projects Can Fail. The Wall Street Journal: CIO Journal. Retrieved (16.3.2021) from: https://deloitte.wsj.com/cio/2012/06/06/5-reasons-analytics-projects-can-fail/

Mahoney, J. T., & Pandian, J. R. (1992). The resource-based view within the conversation of strategic management. Strategic management journal, 13(5), 363-380.

Maity, M., & Dass, M. (2014). Consumer decision-making across modern and traditional channels: E-commerce, m-commerce, in-store. Decision Support Systems, 61, 34-46.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.

Martens, D., & Provost, F. (2011). Pseudo-social network targeting from consumer transaction data. Faculty of Applied Economics, University of Antwerp, Belgium.

Mason, M. (2010, August). Sample size and saturation in PhD studies using qualitative interviews. In Forum qualitative Sozialforschung/Forum: qualitative social research (Vol. 11, No. 3).

Metsämuuronen, J. (2005). Tutkimuksen tekemisen perusteet ihmistieteissä.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: the mediating role of dynamic capabilities and moderating effect of the environment. British Journal of Management, 30(2), 272-298.

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. Information & Management, 57(2), 103169.

Mintzberg, H. (1978). Patterns in strategy formation. Management science, 24(9), 934-948.

Mithas, S., Lee, M. R., Earley, S., Murugesan, S., & Djavanshir, R. (2013). Leveraging big data and business analytics [Guest editors' introduction]. IT professional, 15(6), 18-20.

Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. MIS quarterly, 237-256.

Müller, O., Junglas, I., Brocke, J. V., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: challenges, promises and guidelines. European Journal of Information Systems, 25(4), 289-302.

Oh, W., & Pinsonneault, A. (2007). On the assessment of the strategic value of information technologies: conceptual and analytical approaches. MIS quarterly, 239-265.

Palmatier, R. W., Dant, R. P., & Grewal, D. (2007). A comparative longitudinal analysis of theoretical perspectives of interorganizational relationship performance. Journal of marketing, 71(4), 172-194.

Porter, M. E., & Advantage, C. (1985). Creating and sustaining superior performance. Competitive advantage, 167, 167-206.

Posti. (9.10.2020) Suuri verkkokauppatutkimus 2020. Retrieved 17.02.2021 from: https://www.posti.fi/fi/yrityksille/vinkit-ja-caset/verkkokauppa/suuri-verkkokauppatutkimus-2020

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision-making. Big data, 1(1), 51-59.

Raden, N. (2010). Get Analytics Right from the Start. Sybase, Hired Brains Research.

Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You may not need big data after all. Harvard business review, 91(12), 90-98.

Saddle River, NJ.

Scherpereel, C. M. (2006). Decision orders: a decision taxonomy. Management Decision. 44 (1), 123–136.

Sint, R., Schaffert, S., Stroka, S., & Ferstl, R. (2009). Combining unstructured, fully structured and semi-structured information in semantic wikis. In CEUR Workshop Proceedings, 464, 73-87.

Souza, G. C. (2014). Supply chain analytics. Business Horizons, 57(5), 595-605.

Su, A. J., Hu, Y. C., Kuzmanovic, A., & Koh, C. K. (2014). How to improve your search engine ranking: Myths and reality. ACM Transactions on the Web (TWEB), 8(2), 1-25.

Sumbal, M. S., Tsui, E., & See-to, E. W. (2017). Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector. Journal of Knowledge Management.

Tallon, P. P., Ramirez, R. V., & Short, J. E. (2013). The information artifact in IT governance: toward a theory of information governance. Journal of Management Information Systems, 30(3), 141-178.

Teece, D. J. (2014). The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. Academy of management perspectives, 28(4), 328-352.

Turban, E., Aronson, J. E., & Liang, T. P. (2005). Decision Support Systems and Intelligent System,(Sistem Pendukung Keputusan dan Sistem Cerdas) Ed. 7. Jld. 2.

Vossen, G. (2014). Big data as the new enabler in business and other intelligence. Vietnam Journal of Computer Science, 1(1), 3-14.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. International Journal of Production Economics, 176, 98-110.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, xiii-xxiii.

Williamson, O. E. (1981). The economics of organization: The transaction cost approach. American journal of sociology, 87(3), 548-577.

Wixom, B. H., Yen, B., & Relich, M. (2013). Maximizing value from business analytics. MIS Quarterly Executive, 12(2).

Xiao, X., Tian, Q., & Mao, H. (2020). How the interaction of big data analytics capabilities and digital platform capabilities affects service innovation: A dynamic capabilities view. IEEE Access, 8, 18778-18796.

Yin, R. K. (1994). Case study research - design and methods. Newbury Park, Ca: SAGE.

Zhang, G., Lu, J., & Gao, Y. (2015). Multi-level decision-making. Models, Methods and Applications.

APPENDIX 1: STRUCTURE OF THE INTERVIEW IN ENGLISH AND FINNISH

Introduction

1. Introducing myself
2. Introducing the purpose of the study
3. Informing interviewee about the anonymization and privacy
4. Informing interviewee about the right to not answering a question
5. Informing interviewee about the right to stop the interview
6. Obtain consent to record the interview and use the results in scientific research

Background information of the interviewee

1. Tell me about your career / professional background
    a. What is your professional background?
    b. What is your current job title?
2. Tell me about your daily work
3. How long you have been working in e-commerce?
4. How many employees is in the e-commerce company you are working in?
5. How old is the e-commerce company you are working in?
6. Please describe how data analytics relates to your work?

Data analytics and e-commerce questions

1. Please define data analytics in the context of e-commerce
2. On your opinion, what are the benefits that data analytics provide in e-commerce?
3. On your opinion, what are the weaknesses of data analytics in e-commerce?
4. Do you find the use of data analytics necessary in the context of e-commerce?
5. Do you experience any particular problems with utilizing data analytics?

Data-driven decision-making questions

1. Generally, how can the information that data analytics provide be utilized in organizational decision-making?
2. Tell me about leveraging data analytics in your organization's decision-making
3. Are there any problems or challenges that prevent the use of data analytics in decision-making?
4. Do you think that data analytics should be utilized more in decision-making in your organization?
    a. If yes, what should change so your organization could utilize more from data analytics?

Data analytics capabilities questions

1. Presenting and explaining the table to the interviewee
2. What resources, expertise, and capabilities does your company need in terms of data analytics? and are these resources enough and easy to manage / acquire?
3. Are there any factors / challenges that limit the use of this resource in your organization?
4. Are there any other things that have not yet been mentioned that facilitate or limit the use of data analytics in the organization?

Termination of the interview

1. Asking if there is anything in the answers that the interviewee wants to clarify, change or specify
2. Thanking the interviewee
3. Telling the interviewee how the process will move forward

Johdanto

1. Itseni esittely
2. Tutkimuksen tarkoituksen esittely
3. Ilmoitetaan tulosten anonymisoinnista ja yksityisyyden suojaamisesta
4. Ilmoitetaan haastateltavalle oikeudesta olla vastaamatta kysymykseen
5. Ilmoitetaan oikeudesta keskeyttää tai lopettaa haastattelu
6. Hankitaan suostumus haastattelun nauhoittamiseen ja tulosten käyttämiseen tieteellisessä tutkimuksessa pyytämällä lupa haastateltavalta

Haastateltavan taustatiedot

1. Kerro urastasi / ammatillisesta taustastasi
2. Mikä on nykyinen työnimikkeesi?
3. Kuvaile päivittäistä työtäsi
4. Kuinka kauan olet työskennellyt verkkokaupan parissa?
5. Minkä kokoinen ja minkä ikäinen verkkokauppa on, jonka parissa työskentelet?
6. Kuvaile, kuinka data-analytiikka liittyy työhösi

Data-analytiikka ja verkkokauppa -kysymykset

1. *Määrittele omin sanoin mitä data-analytiikalla tarkoitetaan verkkokaupan yhteydessä*
2. Kuvaile data-analytiikan tarjoamia hyötyjä verkkokaupalle
3. Tuleeko mieleesi mitään heikkouksia tai ongelmia data-analytiikkaan liittyen?
4. Pidätkö data-analytiikan hyödyntämistä välttämättömänä modernille verkkokaupalle?

Data-ohjautuvan päätöksenteon kysymykset

1. Kerro miten data-analytiikkaa vaikuttaa sinun organisaatiosi päätöksentekoon
2. Liittyykö data-analytiikan hyödyntämiseen jotain haasteita tai ongelmia?
3. Tulisiko mielestäsi data-analytiikkaa hyödyntää entistä enemmän päätöksenteossa organisaatiossasi ja mitä tulisi tehdä / muuttaa, jotta data-analytiikasta voitaisiin hyötyä enemmän?

Data-analytiikkakyvykkyys kysymykset

1. Mitä resursseja, osaamista tai kyvykkyyksiä teidän yrityksenne tarvitsee data-analytiikan osalta?
2. Onko näitä resursseja tarpeeksi ja onko niitä helppo hankkia ja hallita?
   a. Tarvitaanko jotain resursseja lisää tai mitä tulisi tehdä, että resursseja voitaisiin hyödyntää tehokkaammin?
3. Tuleeko mieleesi haasteita, jotka rajoittavat jonkun resurssin hankkimista tai käyttöä?

Haastattelun päättäminen

1. Kysytään, onko vastauksissa jotain, jota haastateltava haluaa selventää, muuttaa tai täsmentää
2. Haastateltavan kiittäminen
3. Kerrotaan haastateltavalle, miten tutkimusprosessi tästä etenee.