Eeli Autio

LEVERAGING DATA IN ORGANIZATIONAL DECISION MAKING

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF INFORMATION TECHNOLOGY
2021

TIIVISTELMÄ

Autio, Eeli
Leveraging data in organizational decision making
Jyväskylä: Jyväskylän yliopisto, 2021, 24 s.
Tietojärjestelmätiede, kandidaatintutkielma
Ohjaaja: Taipalus, Toni

Digitalisaation myötä dataa tuotetaan ja on saatavilla enemmän kuin aiemmin. Dataa voidaan hyödyntää organisaatiollisessa päätöksenteossa. Dataohjautuva päätöksenteko vaatii toimia kuten työntekijöiden kouluttamista tai työntekijöiden palkkaamista riittävän osaamisen saavuttamiseksi liittyen dataohjautuviin metodeihin. Dataohjautuvan päätöksenteon implementointi voi muovata organisaation liiketoimintamallia. Tässä kirjallisuuskatsauksessa avataan mitä dataohjautuva päätöksenteko on, kuinka sitä voidaan hyödyntää ja mitä organisaatiolta vaaditaan sen hyödyntämiseksi. Tärkeimmät löydökset olivat, että dataohjautuvan päätöksenteon omaksuminen johtaa parempaan päätöksentekoon ja parantaa organisaation suoriutumista. Dataohjautuva päätöksenteko johtaa myös johdonmukaisempiin päätöksiin. Tutkimuksessa tuli ilmi myös negatiivisia tuloksia kuten algoritmien käyttöön ja yksilöllisen datan keräämiseen liittyvät eettiset ongelmat.

Asiasanat: dataohjautuva päätöksenteko, päätöksenteko, organisaatiollinen päätöksenteko, data-analytiikka

ABSTRACT

Autio, Eeli
Leveraging data in organizational decision making
Jyväskylä: University of Jyväskylä, 2021, 24 pp.
Information Systems, Bachelor's Thesis
Supervisor: Taipalus, Toni

Due to digitalization, there is more data produced and available than before. Data can be leveraged in organizational decision making. Data-driven decision making requires actions such as educating or hiring employees with skills and knowledge about methods and concepts related to data-driven methods. Implementing data-driven decision making into an organization can fundamentally change the organizations business model. This literature review explains what data-driven decision making is, how it can be utilized and what is required from an organization to utilize it. The most important findings were that adoption of data-driven decision making leads to better decision making and enhances the performance of an organization. Data-driven decision making also leads to more consistent decision making. There were also important negative findings such as ethical problems regarding the use of algorithms and gathering of personal data.

Keywords: data-driven decision making, decision making, organizational decision making, data-analytics

TABLES

TABLE OF CONTENTS

1 INTRODUCTION

Due to digitalization, organizations have large amounts of data available for use. There is more data flowing through the internet every second than the entire internet contained 20 years ago (McAfee & Brynjolfsson, 2012). Combined to the development of data storage and processing technologies this has led to new opportunities to leverage data (Brynjolfsson & McElheran, 2016). Organizations can use these opportunities to gain a competitive advantage. As described by Brynjolfsson and McElheran (2016), advancements in IT have made more aspects measurable and analyzable, combined with organizations that heavily invest in IT having even more digital information to draw from. When adding a data-driven culture to this kind of organization, the organizations IT gains a boost that reciprocally leads to better inputs for data-driven decision making. (Brynjolfsson & McElheran, 2016). These possibilities have led to data-driven decision making being appealing for organizations. Generally, there is a notable shift towards basing decisions more on data-based analytics than relying on intuition (Brynjolfsson, Hitt & Kim, 2011).

Important decisions have a significant impact regarding the future. Decisions steer the organization to a certain path that influences the future decisions. Especially in the public sector policies and regulations have far-reaching effects that must be considered carefully (Borison & Hamm, 2010). For decisionmakers it is vital to have a clear picture of what effects the decision has in the future and how it impacts future decision making. Changes in the organization due to employee turnover can lead to inconsistent decision making. Implementing objective decision making methods can lead to more consistent decisions in the long run.

As described by Borison and Hamm (2010), not all strategic decisions are as significant but some strategic decisions that are made regularly can have a significant impact in the future. These decisions affect the stakeholders in the private sector such as employees, customers, suppliers, and stockholders and in the public sector these stakeholders can be citizens. These kinds of decisions have a high level of responsibility and should be made using the best possible information available and their effects on the stakeholders should be acknowledged. (Borison & Hamm, 2010). Stakeholder communication is vital for organizations.

This thesis does not discuss big data and how it is related to this subject. This decision is made to narrow the scope of research to fit its context. The decision is based on the difficulty of defining big data. An attempt to discuss big data would have required searching literature to define big data which would have steered the thesis away from its intent and would have added more content than needed. Some of the used literature do discuss big data and its relation to data-driven decision making but in this thesis I refer only to data when using said literature.

1.1 The research question and the research method

This thesis aims to answer the following research question:

1. How leveraging data in organizational decision making affects the organization?

This study was conducted as a literature review and the answer for the research questions was derived from existing literature. The literature was searched from JYKDOK using search statements formed from the combinations of the following keywords: "data driven decision making", "organizational decision making", "business".

1.2 Research structure

This thesis consists of four chapters. The first chapter is the introduction, which explains the motivation behind the study, the used research method, and the structure of the thesis.

The second chapter defines what data-driven decision making is and discusses how it can be implemented in an organization, what must be taken note of when implementing it, and what is the human role in a data-driven environment. This chapter also discusses the main characteristics of data-driven decision making.

The third chapter presents the results of the literature review. First the benefits of data-driven decision making are discussed in detail and then the possible negative effects are discussed.

The fourth chapter presents the conclusion to the thesis. Motivation behind the study, the utilized research method, the most relevant findings, and the practical implications of the results are summarized. Possible future research possibilities are proposed.

.

2 DATA-DRIVEN DECISION MAKING

This chapter aims to define what data-driven decision making is, what are the characteristics of data-driven decision making, how humans fit into the data-driven environment, and finally how data-driven decision making can be implemented in an organization and what has to be taken note of in this process.

2.1 What is data-driven decision making?

There is a predominant idea of important decisions being made based on experience and intuition of high ranked members of an organization. Accurate information should enable greater use of information in decision making thus leading to better performance (Brynjolfsson et al., 2011) which seems obvious: better and larger amount of knowledge makes it easier to make good decisions. But instead of being satisfied with the knowledge of individuals, organizations can tap into larger pool of knowledge. Data-driven decision making can be described as a practice of basing decisions on the analysis of data rather than on intuition (Provost & Fawcett, 2013). For example, a company can choose a new market segment based on what segments other similar companies target or what seems the best segment based on intuition and previous experiences. The data-driven approach would be basing the decision on analysis of data regarding the company's customer data. The decision can be also a combination of data-driven and intuition-based approaches. In many cases, human judgement has a role even when there is a lot of data available for use in decision making (Borison & Hamm, 2010). For example, in the year 2021 using recent data from consumer behaviour to make decisions regarding what kind of products to bring to the market is problematic. The current state of the world due to the COVID-19 pandemic influences the market and thus human judgement must be used to understand the effects it has regarding the consumer behaviour.

Being data-driven implies that the organization has an understanding of how business decision making is intertwined with an understanding of enterprise data assets (Khatri, 2016). In a data-driven organization the knowledge about what data is and how it can be used should not rest only on shoulders of dedicated data scientists. Organizations, where businesspeople do not understand what the work of data scientists consists of, are in a disadvantageous position, because even though they have all the tools and data they end up making the wrong decisions (Provost & Fawcett, 2013). According to Kiron (2017) every employee should have a certain level of base knowledge about data and processes related to data if an organization aspires to be data-driven. But often many employees do not even know what data refers to: raw numbers, customer activity or social media postings? Without this kind of knowledge, it is nearly impossible to describe what data analytics can offer. (Kiron, 2017). These points make it easier to believe that today the reasons to become more data-driven are mostly not technology related. The main obstacle is the lack of understanding how analytics can be used to improve the organization (Hopkins, Lavalle, Lesser, Shockley & Kruschwitz, 2011). According to the study conducted by Maxwell, Rotz and Garcia (2016) having a common understanding of what it means to use data in decision making throughout the organization is required for building a culture that supports data-driven decision making. If there is not a common understanding within the organization, the use of data-driven decision making can be difficult. In the study it was discovered that perceptions of data-driven decision making differed among the organization staff. It was also discovered that organizations tend to collect data but do not always analyze the collected data. Organizations were positive about using data in decision making but were not confident about the ability of doing so (Maxwell et al. 2016). These kind of issues do steer organizations away from data-driven decision making.

Khatri (2016), Leicht-Deopald et al. (2019) and Power (2014) have proposed that there are different types of analyses that can be used to gain insights to make decision making better. The type of analysis should be chosen based on what kind of information is wanted and what kind of decision it is used to support. The techniques can be categorized by dividing them to

- descriptive/retrospective analyses
- predictive analyses
- prescriptive analyses.

Descriptive/retrospective analyses use historical data to explain what has already happened, how these events affect the current state and are often based on basic statistics. (Khatri, 2016; Leicht-Deobald et al., 2019; Power, 2014). An example of a descriptive/retrospective analysis could be an analysis done to provide a visualization of sales from the last year to understand how different product sales have developed. This information can be used to make decision related to how to proceed with certain products. Products that do not sell as hoped could be pulled from the market or marketing campaigns can be planned to boost the sales of the product. Predictive analyses use historical and/or real-time data to forecast what could happen in the future (Khatri, 2016; Leicht-Deobald et al., 2019; Power, 2014). An example of a predictive analysis could be an analysis done to predict the development of the turnover based on the earlier changes in turnover. This information can be used when planning the expansion of the organization. The third type of analyses, the prescriptive analyses take a step further than the predictive analyses by recommending different actions based on predictions (Khatri, 2016; Leicht-Deobald et al., 2019; Power, 2014). An example of a prescriptive analysis could be an analysis providing different options on where a manufacturer should build the next factory. Prescriptive analysis will also offer insights on what will happen after the decision regarding the location is made. These insights could be for example the future costs or difficulties that the factory could face depending on the country or city it is built in. Additionally, data can be used for optimization, which is using algorithms to choose the best alternatives from a set of solutions based on specific parameters and is often done to achieve maximization or minimization of certain aspect (Khatri, 2016). Optimization can be an automated process done in real-time. For example, optimization can be used to formulate expected return of investment.

2.2 Role of humans in data-driven environment

Despite common misconception, data-driven decision making does not get rid of humans when it comes to making decisions. Data-driven processes are not synonymous with automated processes where human effort is not relevant anymore (Kumar et al., 2013). Davenport and Harris (2005) suggest that when decisions are automated, there should be serious consideration about which decisions should and should not be automated. At the moment the automated decision systems are most useful when making decisions that use electronically available information and have to be made frequently and rapidly. (Davenport & Harris, 2005). Decisions that are made rarely should be left for employees, due to automating these kinds of situations being not worth the hassle and resources. These kinds of decisions could be made when hiring staff for notable and important roles. High ranked personnel who have a great impact on the organization are not hired on a regular basis and are required to have a specific skill set and a great fit to the organizations culture and thus should be hired in a non-automated fashion. Hiring lower ranked personnel for lesser roles where there is a high turnover can be an automated process due to these kinds of employees not being too impactful. Additionally, there is a case to be made that the high ranked positions do not get as many applications than the lower ranked positions.

There are also situations where there is not a lot of data available, it is expensive to obtain or it is not available in a digital form, it is reasonable to use human decision making (McAfee & Brynjolfsson, 2012). The lack of data concerns mostly specific or rare instances. The rarity of the situations consequently means that there has not been as much data generated related to the situation as in other situations. Expensive data can be data that requires advanced tools to gather and handle or data that requires certain expertise. Data that is not in a digital form is hard to leverage. Nondigital data requires an effort to digitalize it. Digitalization of data can be a costly and time consuming task

Data-driven processes need different, more specific human effort. In certain situations, causality can be declared only when making assumptions that go beyond those required for prediction methods, which are not easy to test, and that way require specific domain expertise (Athey, 2017). Often an expert of very specific domain can make assumptions based on their impeccable knowledge gained through experience which are correct and speed up the process. Davenport & Harris (2005) suggest that this kind of human knowledge is also useful in cases where there is not enough data to act on thus making a correct decision through data analysis seemingly less probable than allowing the human manager to make the decision. Besides these kinds of scenarios, managers are needed to review, monitor, and evaluate the decisions due them being responsible of the outcome being correct, legal, and ethical. The managers also are responsible for defining the context and limits for the decision-making process. (Davenport & Harris, 2005). Business leaders with great domain expertise are also needed to do what they in a best-case scenario can do: spot great opportunities, identify possible challenges, embrace, and lift other people, and deal with stakeholders (McAfee & Brynjolfsson, 2012). Especially dealing with stakeholders is a sensitive task that requires human expertise. We are still a long way from automating cross-organizational communication in an effective and meaningful way.

The changing requirements of skills and experience in the job market have become an important point of discussion when talking about the future of job market. Due to already risen levels of automation some jobs are heavily supported or already been taken over by automation. This can be written off as something consistent regarding the technological advancements. For example, the need for horse carriages has been trumped by other ways of transport such as busses, and in the process has made the profession of horse carriage driver almost obsolete but has provided new job titles for people to take on such as bus driver. According to Davenport and Harris (2005) in the context of data-driven decision making the new methods and policies can have a great impact on the staffing requirements, especially when it comes to less skilled and less experienced employees. Automated processes reduce the need for low-skilled or entry-level employees and require the leaders to communicate before the implementation that which decision are automated, and which are not. This is vital for the employment picture. The employees staying in the organization should have a refined picture of the automated processes. (Davenport & Harris, 2005).

Human effort is also important when it comes to data governance. According to Kiron (2017), Rodríguez Rojas, Montenegro and Giménez de Ory (2018) the goals and aims of data-driven decision making should be consistent with other organizational goals and there should be defined indicators to measure and evaluate data-driven decision making. Data governance is also responsible for determining what data matters in the given context, what kind of vocabulary is used to communicate about data, who has access to which data, and how the organization can use the data to create value. (Kiron, 2017; Rodríguez Rojas et al., 2018). Especially the context is important. It is not fair to compare umbrella sales between United Kingdom and Western Sahara. Even though there are powerful algorithms and machine learning methods able to determine weights to different parameters based on how much they matter in the given context, there are different reasons why data governance is beneficial when determining the meaningful data. For those who do not have the access to powerful tools manually picking which data is meaningful is perhaps the only option. Even when the tools are available it may be sensitive to narrow down the pool of data used as an input.

As stated by Kiron (2017), effective information management does require humans guided by rules, norms, and culture that all are fluctuating. Human decision makers are needed to ensure the ethicality and morality of decision making. Algorithms do have set rules, but it is hard to develop algorithms that take into consideration norms and culture.

2.3 Implementation of data-driven decision making

There is an undeniable interest in being data-driven, but often there is a belief that the organization does not have enough data to implement data-driven methods, or they lack the knowledge on how to convert data into actionable insights (Kumar et al., 2013). Being data-driven is often seen as something reserved for large organizations with impressive technological infrastructure and a business model that is heavily based on being data-driven. Small and medium businesses often do not have the advanced technologies and often have less data to act on when comparing to their larger counterparts thus making competing with larger organizations harder in a competitive business environment (Arunachalam & Kumar, 2018). Still data-driven decision making is something that can be used by smaller organizations having limited access to data and who do not base their business model on being data-driven even though implementing a greater focus on data-driven decision making can have an impact on the overall business model (Sleep, Hulland & Gooner, 2019). Partly this is possible thanks to the availability of open-source analytics tools and techniques which can be used by anyone (Arunachalam & Kumar, 2018). Still these factors do have a visible impact on who decides to adopt data-driven decision making. Larger organizations having higher employment and multiple units are significantly more likely to implement data-driven methods in their organization than smaller organizations (Brynjolfsson & McElheran, 2016). In addition to organization size, the environment the organization competes in has an effect. Organizations in a dynamic and competitive environment are more likely to have a greater emphasis on data than organizations in a relatively stable and comfortable business environment (Sleep et al., 2019). Being data-driven in a dynamic environment helps to adjust to the everchanging requirements and trends that the business environment poses. Also, the overall attitude towards IT and the readiness to take on data-driven methods have an impact on adopting data-driven methods. Higher IT investment levels and educated employee base are both correlated with adoption of data-driven decision making (Brynjolfsson & McElheran, 2016). Overall, the education of employees is important in the implementation phase.

Dealing with data-driven methods requires some level of understanding about the subject. Implementing data-driven decision making effectively demands a basic understanding of common statistical terms, data collection, and data analysis (Sherrod, Mckesson & Mumford, 2010). In addition to basic statistic skills there should be employees specialized in key skills related to data handling such as cleaning and organizing large data sets that contain structured and unstructured data (McAfee & Brynjolfsson, 2012). If the employees do not fulfil these requirements, actions should be taken to educate them. As suggested by Sleep et al. (2019) in addition to educating current employees, organizations can hire new employees. New employees can take on new roles that complement data-driven decision making. Organizations tend to approach the new requirements set by data-driven decision making is embed the required skills to the organization through a unit specialized in data-driven methods or in other words an analytical unit. (Sleep et al., 2019). This approach is an easy and concrete way to introduce data-drivency into an organization. It can also help to avoid employees revolting against learning required skills by allowing them to have only a basic understanding about the methods and their effect on the organization.

Educating and hiring people to meet the needs of a data-driven organization is a necessary step but is not enough to ensure a successful shift to data-drivency. According to McAfee and Brynjolfsson (2012) organizations have to change their culture away from basing acts on intuition. Organizations have a habit of pretending to be more data-driven than what they are in reality. In some cases executives have made decisions based on intuition and have come up with "supportive data" after the decision has been made. (McAfee & Brynjolfsson, 2012). This kind of pretending is harmful for the organization. Using resources on only seemingly using data is not efficient and does not grant the real benefits of being data-driven. Trueful adoption of data-driven methods leads to using data in decision making becoming ingrained in the culture of the organization (Sleep et al., 2019).

There should be also an effort to identify what data is needed, is the required data available and if it is not how it can be made available (Sherrod et al., 2010). Without required data, it is hard to make good decisions. Accessible and usable data from different sources should be combined into a single data source and converted into a usable form so it can provide value for the entire organization (Sleep et al., 2019). This especially concerns multi-unit organizations that gather large amounts of data from many different sources. When the data is structured and steered to a one data source it is easier to use in decision making.

This also links to the skills and knowledge of the employees. The lesser the knowledge about data and methods related to it, the simpler the presentation and use of data should be to ensure that it can be used by the entire organization.

Most of the implementation related aspects are related to the individuals within the organization. According to study conducted by Magee, Sammon, Nagle and O'Raghallaigh (2016) the technology of data-driven methods is not the most impactful side of the change, but actually the interaction between individuals and the data provides the biggest impact. The way the data impacts the individual employees determines how successful the data-driven approach is in the organization. Data visualization has a remarkable role in the implementation phase. Visualization help the users of data to see the meaning of data by allowing the users to notice patterns in data. Visualization by itself can inspire change. (Magee et al., 2016). The better the visualization is the easier it is to understand the data. The simplest visualizations such as basic charts can be understood without specific education or knowledge. Still there is a need to educate the employees about basic statistics and other aspect related to working with data, but proper visualization helps to make the data more actionable. According to study conducted by Weiner, Balijepally, Tanniru and Bujnowski (2015) making performance metrics accessible for employees through visualization increases accountability throughout the organization. One way of visualization is the use of dashboards. Dashboards can provide complex information for decision makers effectively. The dashboards can be used to monitor individual and unit performance with ease thus enabling rewarding those performing well and pushing those who are falling short. This allows managers to react to changes in performances and engage in data-driven decision making. In the study it was found that employees started to monitor their own performance, and this had a positive effect to the overall performance of employees. (Weiner et al., 2015). By conveying relevant data in the right way, the data-driven approach can be used at its best potential. When data-driven decision making is being implemented there should be precise planning of what data is presented to whom and how it is presented.

3 EFFECTS OF LEVERAGING DATA IN ORGANIZATIONAL DECISION MAKING

According to Brynjolfsson et al. (2011) organizations gather detailed data from consumers, suppliers, partners, and competitors and additionally from their own regular operations with the use of Resource Planning (ERP), Supply Chain Management (SCM), and Customer Relationship Management (CRM) systems. Organizations have also moved from passive data collection to actively conducting experiments to develop and test different products and services. (Brynjolfsson et al., 2011). Different systems the organizations use provide passive data continuously which can be used to monitor different aspects of the organization. The data can be used to aid organizational decision making. Data-driven decision making provides benefits for an organization but also poses challenges.

3.1 Benefits

The adoption of data-driven methods is heavily motivated by the seeking of benefits they provide. Data-driven methods provide a plethora of benefits to different kinds of organizations operating in different fields.

Use of data and analytics makes organizations operations more efficient and can change their fundamental business models (Kiron, 2017). Study conducted by Brynjolfsson et al. (2011) showed that organizations that adopt data-driven decision making have a 5-6% higher output and productivity than what can be expected based on their other investments and information technology usage. In IT intensive organizations the role of information technology in driving organizational performance is emphasized. Using technologies that enable a wider scale of data gathering or facilitate more efficient distribution of information within an organization can be expected to lower the expenses and to improve performance. (Brynjolfsson et al., 2011). According to Provost and Fawcett (2013) there is statistical evidence that shows that the more data-driven an organization is, the more productive it is. There is also evidence that data-driven decision making is correlated to higher return on assets, return on equity, asset utilization and market value. This correlation relationship seems to be causal. (Provost & Fawcett, 2013). All this supports the assumption that the adoption of data-driven decision making is improves an organizations efficiency, productivity and performance as a whole.

According to Davenport and Harris (2005) data-driven decision making helps to make more consistent decisions than those made by humans and data-driven decision making also hastens the process of turning insights to decisions and decisions to actions. Data-driven decision making has the potential to reduce labour costs, help to leverage scarce expertise, improve quality, and makes responding to customers faster. (Davenport & Harris, 2005). The reduce in labour cost can be achieved through automating simple and repetitive processes which gives employees more time to focus on more important tasks. Scarce expertise can be leveraged by utilizing the expertise in data-driven decision making through enhancing the information to a usable form. Quality can be driven from consistency that being data-driven offers. Having precise and accurate information for the use of decision making should lead to better quality of decisions (Brynjolfsson et al., 2011). Quality can also stem from reducing the judgmental uncertainty in decision making which is caused by individual judgment (Borison & Hamm, 2010). Reducing or removing individual judgment from the decision making process should lead to more objective and consistent decisions. Consistency is vital due to decision made in the present influencing the future decisions. Straying from the chosen path based on individual judgment can lead to unnecessary expenses and loss of time. For example, if an organization wants to widen the field they operate in the decision of where they will steer to will force their hand in the future. A government that takes on a green initiative has to accommodate the fact that they have chosen to be green. In addition to automating the simple and repetitive processes, can also a part of responding to customers be automated. This gives the employees more time to address the important customers. Data-driven methods can also be used to give suggestions on how to respond for customers. For example, an automated response to an email sent by a client can be as simple as only informing the client that their message has been received. These simple interactions are easy to automate and save a lot of time and effort required from employees.

Organizations operating in traditional industries can gain a competitive advantage by exploiting new data sources (Provost & Fawcett, 2013). According to Power (2014), these new data sources can be helpful when identifying trends and customer needs. When combined with new processing technologies and new analyses these data sources can provide more and better decision support for decision makers. (Power, 2014). This leads to data-driven organizations gaining an upper hand when competing in competitive markets and industries. This could also lead to new kind of meta that requires organizations to adopt data-driven methods if they want to compete with others in their industry. With the amounts of data provided by the new data sources and made possible to handle by processing and storage technologies, decision makers can enhance the accuracy of current decision making and also develop new creative ways to make decisions which enables decision making that was not possible earlier (Khatri, 2016). These new ways of using data in decision making enabled by unexpected new data sources are often spotted by experts outside the industry (McAfee & Brynjolfsson, 2012). These opportunities are valuable for organizations that seek to differentiate themselves from their competitors.

According to Hedgebeth (2007) in the current competitive knowledge-based economy organizations need the assistance of different tools to collect, analyse and disseminate information to leverage data in a way that helps knowledge workers to make informed decisions. Fast paced global economy makes the ability to access actionable data valuable. Actionable data provides insights of current performance, customer behaviour and can help to respond to the changing trends. Having accurate and correct data for the use of analytics tools can provide crucial insights to support decision making. (Hedgebeth, 2007). The value of insights based on data is highlighted when trying to predict the future trends. The ability to notice trends at the beginning of their lifecycle can provide a competitive advantage when compared to organizations failing to address the current trends. These kinds of insights can also present new opportunities that the human intuition could not have identified. For example, the possible market segments identified by an organization can fluctuate with time and birth of new trends. Having accurate insights about the development of market segments helps to widen or change the scope of an organization.

Study conducted by Hopkins et al. (2011) found out that organizations that strongly agreed with the idea of using business information and analytics allowing them to differentiate from other organizations within their industry were twice as likely the top performers than the lower performers in their industry. These top performers have a different approach to business operations compared to their peers, especially when it comes to the use of analytics. Top performers tend to use analytics in the widest possible range of decisions. (Hopkins et al., 2011). These findings support the claim of data-driven decision making improving the overall performance of an organizations.

TABLE 1 The main benefits

General benefits	**Benefits for decision making**
Improved efficiency	Consistent decisions due to reduced bias
Improved productivity	Better quality decisions
Lower expenses in some areas	Automation of repetitive decisions
Better overall performance	
Improves competitiveness	

3.2 Challenges

Use of data-driven decision making comes with its own challenges. Some of the challenges are related to the use and implementation of the methods, which must be dealt with for the data-driven decision making to be effective and as correct as possible. Some of the problems are ethical and thus crucial for the organizations deal with. An important challenge that is related to the use of data is bias of used data sets and methods.

3.2.1 Difficulties related to the implementation

It is difficult to standardize formats and establish a standard vocabulary that can be used to facilitate data linking and extending the scope of data-driven decision making to a larger level such as a national level (Rodríguez Rojas et al., 2018). Common language regarding data makes collaborative efforts easier but finding a common language can be a difficult and time consuming task. In addition to a common language regarding methods and basic concepts, the datasets should be aligned with each other. Categorizing data also poses a challenge for organizations (Kiron, 2017). A government could implement a taxing policy regarding roads. To implement such a policy, there should be categories for different types of roads that require certain tax to use. This would require a definition of a road and its subcategories. It should also be noted that the same definitions should be used elsewhere in the organization: using different definition of road in the transport and tax contexts could cause problems.

In addition to the struggle of establishing a standard vocabulary, there are often gaps in data quality (Kiron, 2017). These gaps can be lack of data related to a certain group. An example of a data gap could be an organizations marketing unit missing sales data categorized by age group. The unit probably has data of sales as a whole but it is not categorized by age group which leads to having to take other measures to fill the gap such as surveys aimed for different age groups. Recognizing data gaps and filling them improves the quality of data-driven decision making but requires research and work to achieve.

Unstructured data also poses challenges when implementing data-driven decision making (Kiron, 2017). As described by Arunachalam and Kumar (2018) different kinds of log files gathered from organizations systems can differ in their formatting and be unstructured. In most cases unstructured data must be structured before utilizing. Having only structured data does not get rid of all the problems. Certain types of data such as Likert-scale type data are difficult to utilize in certain contexts. For example, the Likert-scale type data often gathered actively with questionnaires and other tools does not reflect the intensity between different choices available for the responder. These kinds of situations can be solved by using suitable distance measures that indicate the intensity. (Arunachalam & Kumar, 2018).

3.2.2 Ethical problems

Use of data often presents legal and ethical problems. Legislation and ethicality can set borders on how much individual data can be gathered and used in decision making thus limiting the use of automated decision processes in certain contexts (Davenport & Harris, 2005). Legislation such as GDPR limit the gathering and use of individual data. Organizations benefit from having more data at their disposal, but this requires responsibility regarding individual privacy. According to Kumar et al. (2013) this entails ensuring data privacy and security at all levels of data collection. Organizations must obtain certain permissions from individuals to collect and store their data. Large scale of individual data collection and security threats it includes can discourage consumers from sharing their personal information. (Kumar et al., 2013). Consumers have become more aware of the issues related to data collection and use of their data. This has led to consumers becoming cautious to some extent when it comes to giving up their data. Still there is a phenomenon of consumers being inconsistent when deciding whether to reserve or give their data. For example, some consumers might not care about some webstore tracking their actions on the website but if they were asked whether it is appropriate if a physical store tracks their purchases, they might not be so supportive towards data collection. In recent times the discussion about privacy and the ethicality of collecting data from consumers has been on the rise and warrants actions from organizations to ensure that their processes are acceptable. But still the whole discussion is a grey area: most people are aware of the problems but do not care enough to change their actions or give up on the benefits they receive when an online store tracks their actions or when a music application suggests music to them based on what they have listened for before.

In addition to privacy concerns there are concerns about inequality. Using algorithms in recruitment can lead to inequality based on gender or race (Athey, 2017; Leicht-Deobald et al., 2019). The bias can stem from either the used algorithms, datasets, or both. According to Leicht-Deobald et al. (2019) inequality issues can arise in HR processes such as recruiting, employee evalution and performance appraisals. Using algorithms in these processes can be effective but also ethically problematic. Algorithms trained using historical data can have a bias towards certain groups. (Leicht-Deobald et al., 2019). For example, in fields where there is a male dominance the algorithms can be biased towards men. The algorithm can learn to be biased through analysing historical data: there is a combining factor of gender amongst the previously recruited employees. This can be coincidental due to males choosing the profession more often than other genders tend to thus leading to the majority of applicants in the field being male. This can lead to discrimination based on gender. To avoid these kinds of problems the used datasets should be catered to not containing information that can impair the recruitment process thus leading to objective outcomes.

There is also a growing concern towards the lack of transparency of the algorithms and the question of who is to be held accountable (Leicht-Deobald et al., 2019). The lack of transparency invites questions about the objectivity and the fairness of decisions. Using data-driven decision making to decide who is hired or who is granted income support affect the individuals whom the decisions concern. These individuals should be provided with reasonings on why such a decision was made. Algorithms making the decisions can be hard to visualize and understand without required knowledge of algorithmics. This leads to most algorithms being black boxes to the public. To be ethical the decision making algorithms should be transparent to a certain point. The question of who is accountable should also be thought of when leaning towards data-driven decision making. If data-driven decision making is used to determine what kind of healthcare a certain individual receives, who is responsible of the individual if the decisions lead to negative outcomes? This question poses a challenge for the implementation of data-driven methods to sensible areas.

TABLE 2 The main challenges

Ethical challenges	**Challenges related to implementation**
Possible inequality due to biased datasets	Hard to standardize formats
Privacy concerns	Categorizing datasets is hard
Lack of transparency	Gaps in data can affect the outcomes
Legislation related to privacy	Structuring and formatting data is hard

4 CONCLUSIONS

This thesis was a literature review which discussed leveraging data in organizational decision making. The main focus was on data-driven decision making. The scope of the research was narrowed to not include big data due to it being such a complex topic which would have required focusing too much on defining what is big data. This thesis also provided examples in some contexts but as a whole is aimed to give a general view of leveraging data in organizations. This thesis aimed to answer the research question "How leveraging data in organizational decision making affects the organization?"

The first chapter was the introduction, which explained the motivation behind the study, the used research method, and the structure of the method. The growing amount of data that is available for organizations combined with the development of better tools and technologies for the storing and handling of data has led to new opportunities to leverage data.

The second chapter defined what data-driven decision making is and discussed its implementation to an organization. The human role in a data-driven environment was also discussed. Data-driven decision making can be described as a practice of basing decisions on the analysis of data rather than on intuition (Provost & Fawcett, 2013). Data is collected actively with surveys and other methods or passively from different systems that the organization uses. The gathered data is processed with the help of analytics tools and transformed to insights that can be used to aid the decision making process.

Humans have their own role to play in a data-driven environment. Humans are needed to assess which decisions should be automated and which should not be automated (Davenport & Harris, 2005). In situations were there is not a lot of data available, or it is not sensible to obtain decision are wiser to be made by humans. This is the case also when the situation is rare, and the related decisions are not made often. Humans are also needed to judge the ethicality of a decision and how well it is in line with the organization's goals.

Effective implementation of data-driven decision making demands a basic understanding of common statistical terms, data collection, and data analysis (Sherrod et al., 2010). If an organization lacks these qualities, actions should be taken to obtain a satisfying level of knowledge. Organizations can educate their existing employees, or they can hire new employees who possess the needed knowledge and skills (Sleep et al., 2019). In addition to sufficient knowledge and skills, an organization needs to identify which data is needed, is it available and if not, how it can be made available (Sherrod et al., 2010). Without sufficient amount of accurate data, data-driven decision making can not be leveraged to its fullest.

The third chapter presented the results of the study. The benefits were first discussed, and the main benefits can be found from the table 1. Then the negative effects were discussed, and the main negative effects can be found from the table 2. The main benefits that were found are improved efficiency, productivity, competitiveness, overall performance, decision consistency, decision quality, automation of repetitive decisions and lower expenses in some areas. The main challenges that were found are possible inequality due to biased datasets, privacy concerns, lack of transparency, legislation related to privacy, standardizing formats, categorizing datasets, gaps in data, and structuring and formatting data.

Data-driven decision making provides a plethora of benefits for an organization. The use of data is a growing trend and its use to aid decision makers is even more relevant in the future due to constant development of better data storage and processing technologies. Also, it can be argued that in the future the education of employees provides them with better knowledge and skills to deal with information technology. This thesis provides value for those responsible of making decisions. The information provided can be used by decision makers in many fields to understand what the requirements are for implementing data-driven methods, what benefits they can expect and what challenges they must address. Mainly the decision makers can get a basic level of understanding about the topic which allows them to invest themselves further into the subject. Future research could focus more on how small and medium organizations could be encouraged to adopt data-driven decision making. An interesting research topic could also be exploring different levels of implementation of data-driven decision making. Proposing models of different levels of implementation could encourage more organizations to adopt data-driven decision making and would make data-drivency more accessible for different kinds of organizations. Case studies that study specific organizations operating in different markets and fields could also provide valuable insights about the implementation of data-driven decision making and the benefits and challenges it poses

REFERENCES

Arunachalam, D. & Kumar, N. (2018). Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Systems with Applications*, 111, 11.

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science, 355*(6324), 483-485.

Borison, A. & Hamm, G. (2010). Prediction Markets: A New Tool for Strategic Decision Making. *California Management Review*, 52(4), 125-141.

Brynjolfsson, E., Hitt, L. M. & Kim, H. H. (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? *SSRN Electronic Journal*.

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *The American Economic Review, 106*(5), 133-139.

Davenport, T. H., & Harris, J. G. (2005). Automated decision making comes of age. *MIT Sloan Management Review, 46*(4), 83.

Hedgebeth, D. (2007). Data-driven decision making for the enterprise: An overview of business intelligence applications. *VINE, 37*(4), 414-420.

Hopkins, M. S., Lavalle, S., Lesser, E., Shockley, R. & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value.(The New Intelligent Enterprise)(Report). *MIT Sloan Management Review, 52*(2), 21.

Khatri, V. (2016). Managerial work in the realm of the digital universe: The role of the data triad. *Business Horizons, 59*(6), 673.

Kiron, D. (2017). Lessons from becoming a data-driven organization. *MIT Sloan Management Review, 58*(2), 0.

Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A. & Henseler, J. (2013). Data-driven services marketing in a connected world. *Journal of Service Management, 24*(3), 330-352.

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I. & Kasper, G. (2019). The Challenges of Algorithm-Based HR Decision-Making for Personal Integrity. *Journal of business ethics : JBE, 160*(2), 377.

Magee, B., Sammon, D., Nagle, T. & O'Raghallaigh, P. (2016). Introducing data driven practices into sales environments: Examining the impact of data visualisation on user engagement and sales results. *Journal of Decision Systems, 25*(sup1), 313-328.

Maxwell, N. L., Rotz, D. & Garcia, C. (2016). Data and Decision Making: Same Organization, Different Perceptions; Different Organizations, Different Perceptions. *The American journal of evaluation, 37*(4), 463.

McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review, 90*(10), 61-68.

Power, D. J. (2014). Using 'Big Data' for analytics and decision support. *Journal of decision systems, 23*(2), 222-228.

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

Rodríguez Rojas, L., Montenegro, C. & Giménez de Ory, E. (2018). Metamodel to support decision-making from open government data. *Journal of Ambient Intelligence and Humanized Computing, 9*(3), 553-563.

Sherrod, D., Mckesson, T. & Mumford, M. (2010). Are you prepared for data-driven decision making? *Nursing management, 41*(5), 51.

Sleep, S., Hulland, J. & Gooner, R. (2019). THE DATA HIERARCHY: Factors influencing the adoption and implementation of data-driven decision making. *AMS Review, 9*(3), 230-248.

Weiner, J., Balijepally, V., Tanniru, M. & Bujnowski, A. (2015). Integrating Strategic and Operational Decision Making Using Data-Driven Dashboards: The Case of St. Joseph Mercy Oakland Hospital/PRACTITIONER APPLICATION. *Journal of Healthcare Management, 60*(5), 319-331.