LUT UNIVERSITY
School of Business and Management
Master's Degree Program in Strategic Finance and Business Analytics

Master's thesis

Automation of BI processes for reporting and data visualisation: Case OP Markets

Janne Kainulainen

17.5.2020

Supervisor and 1st Examiner: Mikael Collan

Second Examiner: Jyrki Savolainen

ABSTRACT

Author:	Janne Kainulainen
Title:	Automation of BI processes for reporting and data visualisation: Case OP Markets
Faculty:	LUT School of Business and Management
Degree:	Master of Science in Economics and Business Administration
Master's Programme:	Strategic Finance and Business Analytics
Year:	2020
Master's Thesis:	LUT University 67 pages, 14 figures, 7 tables
Examiners:	Professor Mikael Collan Research Fellow Jyrki Savolainen
Keywords:	Business Intelligence, Business Analytics, BI process, BI architecture, automation, challenges

Business Intelligence and Analytics (BI&A) as an industry has experienced rapid growth and attracted a lot of academic interest in the past few decades. Interest towards closer to real-time decision-making based on data has increased in both academic and business fields. This work aims to build an in-depth understanding between the theory and practice of automated Business Intelligence (BI) processes. As an outcome, it is expected that the findings from this study can improve the current BI implementation of the observed case company. This study is conducted as a qualitative research. A theoretical framework is created to build a comprehensive view of the term BI. Literature review section provides an overview on the academic literature about BI use cases and faced problems in the implementation. In the case study implementation of BI process automation and found problems in a Finnish bank are observed.

According to the study, practice can benefit from the theory and vice versa. Problems faced in the earlier studies seems to be similar to the observations made in the case study. It seems that organisations deal with multiple challenges simultaneously rather than just with one specific problem. The established theoretical framework can be utilized to evaluate BI implementation and make improvements based on the findings. In addition to this, the general BI process model proposed in this study seems to be applicable in evaluating the BI process implementation in practice.

TIIVISTELMÄ

Tekijä:	Janne Kainulainen
Tutkielman nimi:	BI prosessien automaatio raportointia ja datan visualisointia varten: Case OP Markets
Tiedekunta:	LUT-kauppakorkeakoulu
Tutkinto:	Kauppatieteiden maisteri
Maisteriohjelma:	Strateginen rahoitus ja bisnesanalytiikka
Vuosi:	2020
Pro Gradu -tutkielma:	LUT-yliopisto 67 sivua, 14 kuvaa, 7 taulukkoa
Tarkastajat:	Professori Mikael Collan Tutkijaprofessori Jyrki Savolainen
Avainsanat:	Business Intelligence, Business Analytics, BI prosessi, BI arkkitehtuuri, automaatio, haasteet

Business Intelligence ja Analytiikka (BI&A) toimialana on kasvanut huomattavasti ja herättänyt paljon akateemista huomiota viimeisten vuosikymmenten aikana. Kiinnostus reaaliaikaista dataan pohjautuvaa päätöksentekoa kohtaan on kasvanut niin akatemian kuin liiketoiminnan osa-alueilla. Tämän tutkielman tavoitteena on rakentaa syvällistä ymmärrystä teorian ja käytännön välillä automatisoitujen Business Intelligence (BI) prosessien osalta. Tutkimuksen tulosten avulla odotetaan saatavan kehitysideoita case yrityksen BI prosessin nykyiseen toteutukseen. Tämä tutkimus on toteutettu laadullisena tutkimuksena. Teoreettisen viitekehyksen avulla luodaan kokonaisvaltainen katsaus BI termistä ja sen käytöstä. Kirjallisuuskatsauksessa tarkastellaan akateemisessa kirjallisuudessa esitettyjä BI:n käyttökohteita ja toteutuksessa kohdattuja ongelmia. Tapaustutkimuksessa tarkastellaan automatisoidun BI prosessin toteutusta ja havaittuja ongelmia suomalaisessa pankissa.

Tutkimuksen mukaan teoriasta voidaan saada hyötyjä käytäntöön ja vastaavasti käytännöt voivat hyödyttää teoriaa. Aiemmissa tutkimuksissa kohdatut ongelmat vaikuttavat olevan samankaltaisia havaittujen ongelmien kanssa. Vaikuttaa siltä, että organisaatiot kohtaavat useita haasteita samanaikaisesti vain yksittäisten ongelmien sijasta. Vakiintunutta teoreettista viitekehystä voidaan hyödyntää BI prosessien toteutuksen arvioinnissa ja parannuksia voidaan tehdä näiden havaintojen perusteella. Tämän lisäksi tutkimuksessa ehdotettu yleistetty BI prosessimalli näyttää sopivan BI prosessien tarkasteluun käytännössä.

Acknowledgements

One chapter in my life is coming to its end. By writing this acknowledgement I'm wrapping up my academic journey (at least for now) that started six years prior with an approval of admission into a university. I must admit that I feel huge relief at the moment when I can conclude my master's studies and move towards new challenges in life. But at the same time, I'm already missing all the student life and sense of community it has provided. It's been a privilege to get known and work with so many great fellow students. We had a good time.

I want to express my gratitude for the teachers and staff of LUT for providing us a high-quality teaching and building an inspiring environment to study. A special thanks goes to my thesis supervisors Mikael Collan and Jyrki Savolainen. I'm grateful for your support and valuable guidance throughout the thesis process. Also, I want to thank the representatives of the case company enabling this study. Especially D.Sc. Juho Lauri for sparring the ideas and helping along the process.

And finally, I am grateful for the unconditional support from my family. Always standing behind me nevertheless the path I have chosen to go along. And the way that you have given me a desire for knowledge makes me always to explore further. Learning is a lifelong journey.

Table of contents

List of figures

List of tables

Abbreviations

BA	Business Analytics
BD	Big Data
BI	Business Intelligence
BI&A	Business Intelligence and Analytics
CI	Competitive Intelligence
CRM	Customer Relationship Management
DDD	Data driven decision-making
DSS	Decision Support Systems
DW	Data Warehouse
ETL	Extract Transform and Load
ERP	Enterprise Resources Planning
IT	Information Technology
MI	Market Intelligence
OLAP	Online analytical processing
SCM	Supply Chain Management

1. Introduction

Business Intelligence and Analytics (BI&A) as an industry is in rapid growth and the business logic of some companies leans completely on data and analytics (Baryak, 2015, pp. 230-231; Davenport & Dyché, 2013, p. 1). Companies have been investing vast amounts of money and resources into business intelligence (BI) tools (Davenport & Dyché, 2013, p. 19) with no slowdown into the foreseeable future. Reasons for these investments come from proven results that BI&A applications can provide. Indeed, notable driver for investing in BI&A tools comes from the urge to gain competitive edge in their own industry and adapt to the changes that digitalisation has caused in the surrounding environment. In addition to these reasons, companies are moving more towards data-driven decision-making (DDD). (Božič & Vlado, 2019, p. 94; Nordic Institute of Business & Society NIBS, 2020, p. 16)

Relying more on BI in their decision-making and planning processes allows companies to harness data to make more informed decisions and try to make predictions about future outcomes. As such, decision makers are able to get a holistic view of their business and customers. This enables them to make data-driven decisions, improve operational efficiency and deliver business-critical solutions. (Baryak, 2015, p. 231) Consider the following real-life business application of BI&A from the retail industry. The New York Times made an interview-based article about how Walmart used data-driven analytics to identify and predict unusual local demand for products ahead of a hurricane. Besides the obvious demand spikes in bottled water and canned food more careful data analysis offered insights about demand for products that were less obvious. In fact, new findings from the data were that strawberry Pop-Tarts increased in sales seven-fold and beer was top-selling item pre-hurricane. (Hays, 2004) To be able to conduct this kind of analysis BI is utilized to treat, compile and manipulate data and present result in a human-understandable form. These kinds of insights are valuable information for businesses enabling data-driven decision making. According to (Kim & Bui, 2019, p. 2) Walmart has been recognized for having one of the best practices in the retail industry for disaster supply chain management, demonstrating the power of appropriately applied BI practices. In the retail industry it is usual to work with the supply chain optimisation and predicting sales also outside of disaster

situations. For example, the Finnish cooperative retail company S-Group has identified data as an excellent source for offering information about customers. This helps them to optimise their supply chain and make predictions about upcoming sales figures. In fact, they see that BI&A is the only way to retain the market share and be able to compete against rivals in the industry. (SAS Institute, 2018)

Nowadays companies maintain more data from internal and external sources than ever. Further, the variety of information types and information sources have also increased greatly. This leads to the problem where companies have to manage and maintain these large masses of data. Unfortunately, traditional database applications do not have the capability to analyse such amounts data and draw conclusions from it for decision-makers. In addition to this it may be challenging to distinguish relevant information from the large data mass. Establishment of the term Big Data (BD) originates from this development indicating the complexity and size of data sets. (Baryak, 2015, p. 231) BI is closely related to BD where it acts as a solution and a method for decision-makers to analyse these data masses. It is difficult to find any successful enterprise that has not leveraged the benefits of BI in its business. (Chauldri, et al., 2011, p. 88) When companies want to make informed decisions based on information and react to changes rapidly it is essential to have the latest possible data available. In the previous decades data for these BI applications have been typically gathered periodically by marketing and finance departments in a form of monthly and yearly reports. Questions and insights that differentiate from the reported form have required ad hoc work and compiling data from different sources. (Azvine, et al., 2005, p. 214) Technical advancements have made possible to automate these processes and make BI closer to real time analytics.

In this research, a case study is performed observing the automation of BI processes in one specific department of a Finnish cooperative OP Financial Group. Before examining the case, a general background on BI processes and how they are generally automated at the moment are given. In particular, after setting the theoretical stage, it is contrasted with the practical setting of the case study, seeing how theory differs from practice and how the findings compare to other studies. It will be interesting to examine what kind of difficulties companies have faced in BI process utilisation and how it

compares to the findings. Motivation for this study is two-fold: to make automated BI processes more achievable for businesses and to potentially suggest or modify the current best practices. This topic is interesting to study because advances in technology, especially in the field of Information Technology (IT), makes automated BI processes more achievable for businesses. Even though studies have shown that top tier companies leverage the benefits of BI it is always possible to strive for better. Automated BI processes using closer to real-time data can be used for data-driven decision-making in the company and help to gain competitive edge against peers in the industry.

1.1. Purpose of the study

This work aims to build an in-depth understanding between theory and practice of automated BI processes. This study has its focus on the automation of BI processes for reporting and data visualization. In particular, the aim of this study is to increase knowledge about automated BI processes and its real-life applications. As an outcome, it is expected that the findings from this study can help on improve the current BI implementation of the observed case company.

As mentioned in the introduction, BI&A as an industry is still growing rapidly and has attracted considerable academic attention as well. As such, it is interesting to study how much the practice actually appears guided by theory. As a first step, the study starts with having a look at the various definitions of the relevant terms via a literature review. As a by-product, the study strives to uncover the aspects of theory that could use further academic research. Moreover, the implementation of BI&A has been seen as an interesting topic because this stage is complex and critical for the creation of successful BI&A application (Côrte-Real, et al., 2014, pp. 174-175). By going through these different definitions this study tries to clarify their meaning and form a comprehensive image about the topic. This is an excellent opportunity to study this subject more in depth and gain more ideas and expertise for the future.

By conducting a case study and making observations it is possible to gain ideas to improve the automated BI processes. On the side it is possible to get new insights about the process that has not been considered before. These ideas and insights can

be turned into actions to improve the process and make it better. This study can bring benefits both for academic and business point of view. From the academic point of view, it is possible to get more clarification on conceptual definitions as discussed above and what kind of BI processes has been automated at the time of writing this study. On the business side perceptions can be made if current automation implementation has been built in a remarkably different way than it is described in theory. Also, going through the current process it can bring up new ideas for improvement and make the process more efficient. It is also possible to find completely new or different ways to implement and utilize BI in the business.

1.2. Research focus

The main idea of this study is to observe how automation of BI processes can be implemented in a real-world setting. BI has been widely discussed in academia and there are many differentiating key research themes under this topic. This study will focus on specific themes in this research to make it a suitable topic for master's thesis. To find these pinpoints for the study it is crucial to start with a working definition for the term BI. Afterwards, it is possible to take a view at the underlying factors that must be considered in order to build an automated BI process. Followed with proceeding to observe real-world use cases of BI. Starting with an illustration of how different scientific fields are closely related to the topic of this study are shown in Figure 1 below.

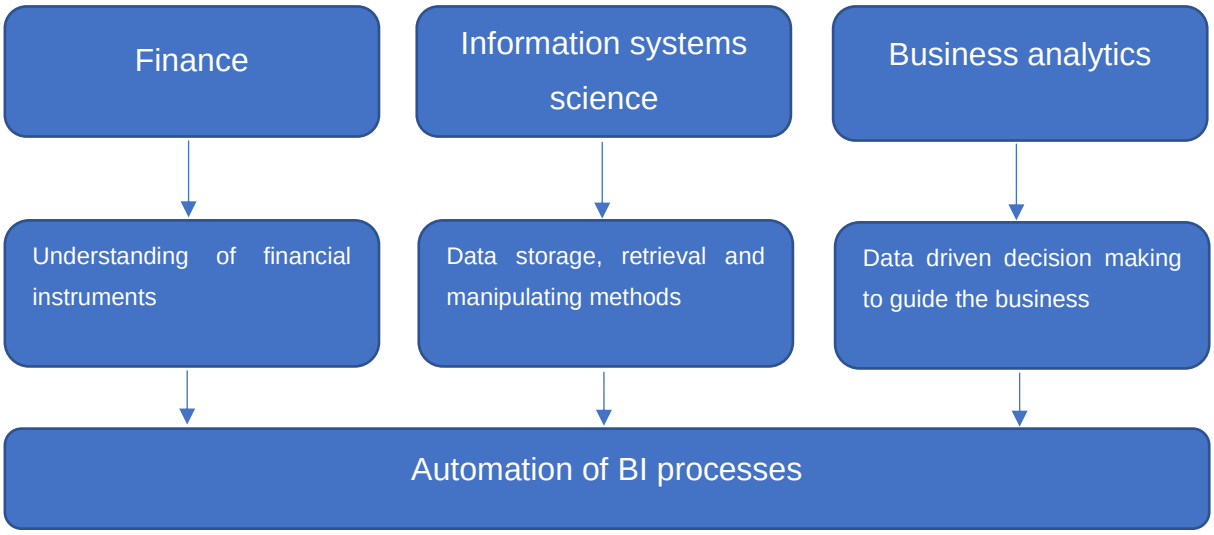

Figure 1 How different scientific field are related to the automation of BI process in this study context

These above-mentioned fields of science contribute to this topic and are related to the theoretical background that is begun with. When this underlying theoretical framework has been established it is reasonable to proceed with a literature review. The literature review section goes through what kind of BI processes has been implemented at the time of writing this study. Examples from earlier studies of BI processes implementation will be examined also. It gives an answer for what kind of BI processes has been automated and what kind of challenges are met in the process. This helps to understand how this specific topic has been already studied and reflect the findings compared to earlier research. After establishing the underlying theoretical framework and conducting a literature review concerning the studied topic it can be followed with an observative case study. With this method it is possible to make observations how BI process automation has been implemented in a unique real-life business setting. Observations and comparison are made to the theoretical framework and earlier research can ramp up more innovative perspective for using BI and its current implementation. Based on the case study results improvement suggestions are given. This will address how the results of the study can be utilized to improve current state of BI in the organisation.

When the steps mentioned in previous paragraph are pulled together the research focus for this study ca be defined with four steps. These steps are presented in the following figure. Writing down these steps helps to narrow down the studied subject and concentrate on a specific topic. The first step in this study is to look at the research area. The research area holds four important aspects that are: BI as a definition itself, banking industry where this study is situated in, automation of the BI processes and data driven decision-making tied to the benefits from BI. These four areas form the base for this study. The second step holds the definitions of research object. The objective of this study is to observe how automation of BI process can be implemented in a real-world setting. Background and requirements for BI process automation were defined and earlier studies about implementation was explored. This gave a baseline for comparing the observations and implementation challenges found in the case study. The third step was to find the perspective that this study addresses the topic. An observative case study method is used in this study where the automation of BI process in a Finnish bank is observed. Therefore, the perspective for this study comes

from the user perspective. This leads to the fourth and last step which is research focus. The focus is on observing challenges met in BI process automation and finding ways from earlier research to improve the studied case company's process.

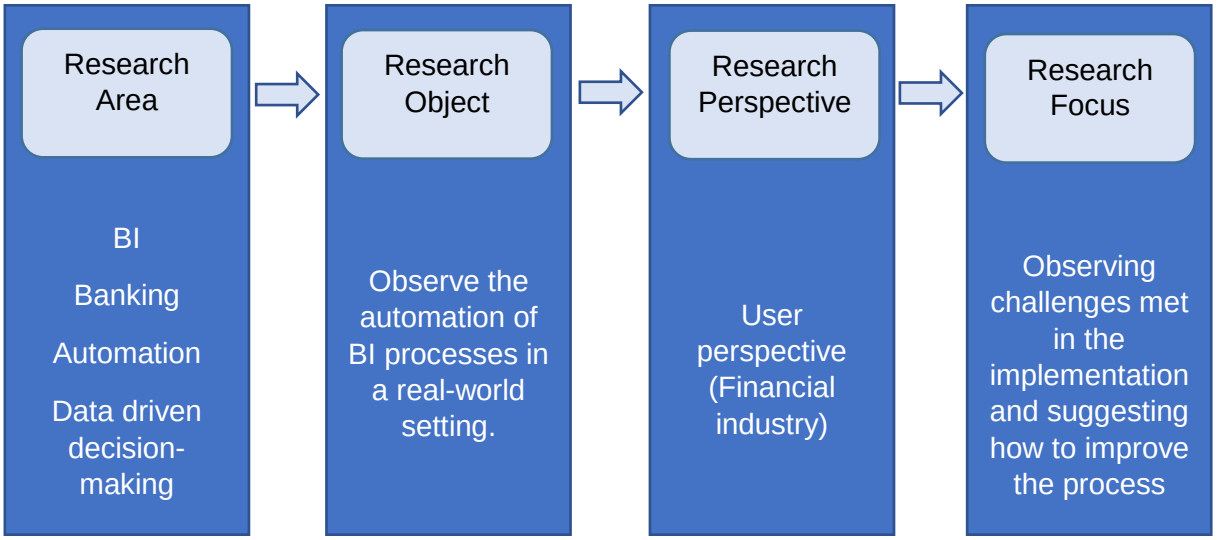

Figure 2 Research focus

1.3. Research questions

In the previous paragraph the focus of this master's thesis is addressed and it leads to one main research question and three supporting sub questions. These three supporting sub questions assist in finding an answer to the main research question.

Firstly, the main research question is: **What kind of challenges are faced in BI process automation in the context of banking business?** This emphasises the main idea of this study to observe differences between the theory and real-life applications of BI processes. An answer to the main research question can be found with the literature review and from the case study observations. The following sub questions support the structure of this study and will gradually lead to the main research question. The first sub question takes the initial step towards answering the main research question by addressing the background:

1. What has been written about BI in the academic literature?

After this, the study moves towards looking for earlier research about typical use cases for BI automation. This will also help identifying topical challenges with the automation of BI processes. The second sub question stands as follows:

2. What kind of BI processes are typically automated and what kind of problems are met in the process?

Third and last sub question is derived from observations made in the real-world context within the focal area of this work presented in theoretical framework and findings in the literature review. This ties the main research question to the research focus presented in Figure 2 of previous subsection. It will also help in forming suggestions of possible improvements for the observed business. Third and last sub question can be formed as follows:

3. How can the findings compared to earlier research help on improving BI process automation?

Information sources for this study are mainly from academic literature where answers for the literature review and background part of the study can be acquired. Professional literature can give insights regarding the process implementation and can be utilized in the background part of the study and to understanding the methods that the company have chosen in its implementation. Observations made in the case study come from company's internal and external sources. External sources are used for making a presentation of the company and internal sources such as documentation and process observations for the case study section.

All these mentioned research questions and information sources are used to gain more knowledge about BI as a concept and the real-world use cases. The background section provides a comprehensive overview about BI and finest details are ruled out from this study. The empirical part which is conducted in form of a case study observes challenges met in the implementation and compares them to the earlier findings that have come across in the literature review section. As an outcome suggestion for improvements in the case company BI implementation can be presented.

1.4. Structure of the study

This master's thesis consists of four parts. The first part includes an introduction which covers an overview of the thesis and its scope. The second part of this study includes a theoretical framework. Starting with an assessment of the background of the topic and continued with literature review. The third part of this thesis includes an explanation of used research method and a case study. The fourth and last part of this thesis include conclusions and discussion about the results. The following figure illustrates the structure of this thesis more closely.

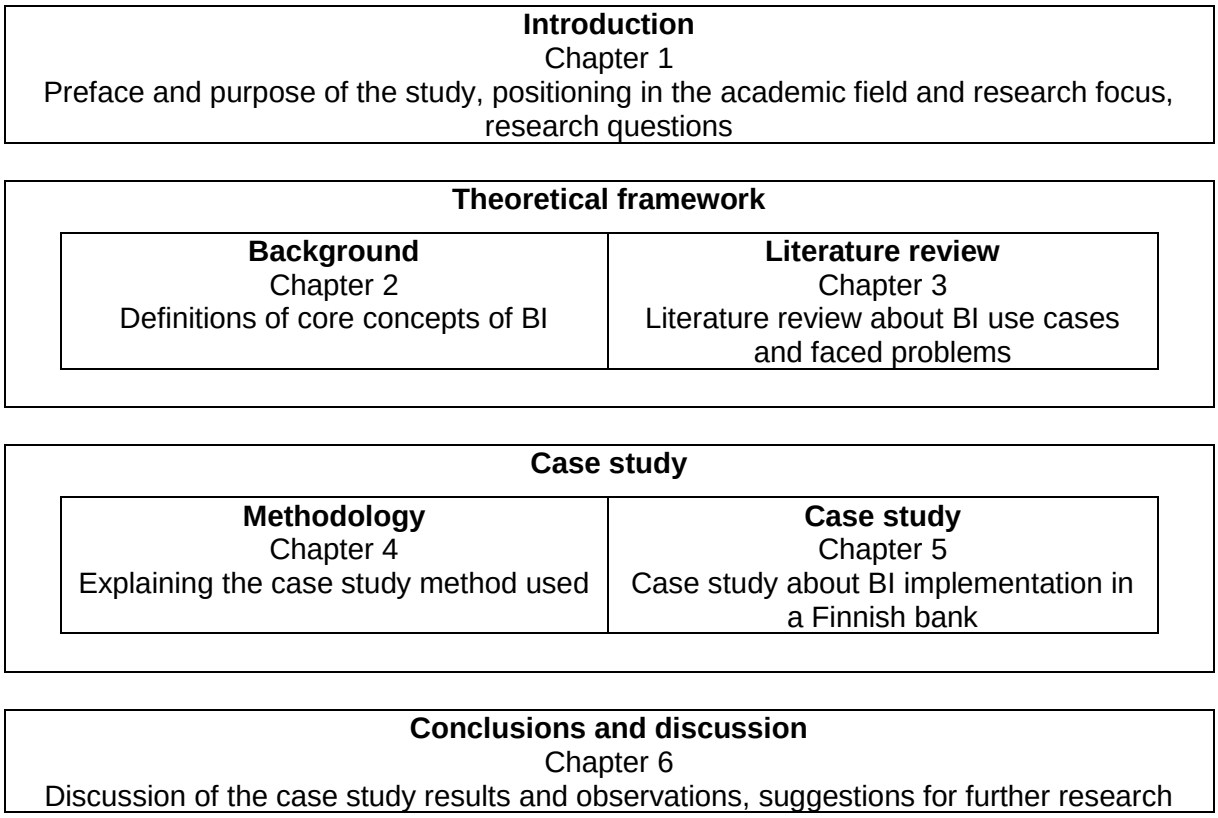

Figure 3 Structure of the thesis

In Chapter 1, the topic and purpose of this study are presented. This leads to defining the research focus which clarifies the scope of this study. Also, research questions and origins for these questions are presented. The second part of this study includes a theoretical framework. In Chapter 2, core concepts of BI are defined to build up a background for the thesis. This is followed by Chapter 3 where a literature review is conducted to find BI use cases in real-world setting. The third part of this thesis

includes the case study. Chapter 4 is denoted for explaining the scientific methods used in the case study. Chapter 5 includes the case study where the implementation of BI process automation is observed in a Finnish bank. Fourth and last part of this thesis consist of Chapter 6, which holds the conclusions and discussion about the results.

2. Background

2.1. A brief history of BI

Using data to improve operational performance and making decisions based on information is not a new innovation in business. Organisations have utilized data for over 200 years. For an example, in the late 19th century companies such as Carnegie Steel and Standard Oil used data to make more informed decisions and improve their business performance. These actions were based on very basic information about their sales and production figures which were used to guide the decision-making in business. (Tedlow, 2003) Despite the early application of the concept, a first definition for the term BI has its roots in the 1950's. After the first introduction of this term it has been defined in different ways during the decades. Hans Peter Luhn defined and stated the objects of Business Intelligence System as: *"the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal"* and also notable is that the main object of such system is to provide suitable information to support actions (Luhn, 1958, pp. 314-315). From those days BI has developed through many phases. Technological advancements and capabilities led to increase in amounts of stored and processed data around 1980's. These advancements and changes in Information Technology (IT) particularised the use of term BI to describing different kind of reporting, decision-making systems and Data Warehouse (DW) solutions and methods. This process transforms raw data into relevant and usable form which opens possibilities for intelligence and strategic knowledge about the business. (Tyson, 1986, p. 9; Hovi, et al., 2009, p. 79) These requirements create an opening for a systemised process for collecting information from different sources and store it in unified format. BI process can then utilize can combine this data in a robust way as shown in Figure 4 below.

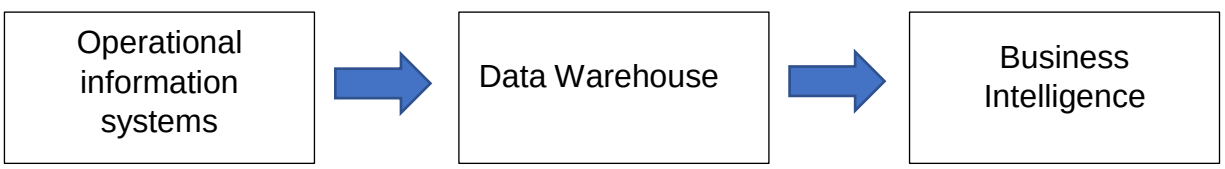

Figure 4 Relation between Data Warehousing and BI after (Hovi, et al., 2009, p. XIII)

In order to use BI, the underlying data needs to be structured and stored somehow. Even though BI solutions can be built without using a DW for compiling and storing data, these solutions have multiple shortcomings. For instance, one could gather information to a spreadsheet file where data is easily editable and shareable. Arguably, many users are familiar with this type of approach. However, this method is prone to errors in data compiling and preparation. Furthermore, it is hard to keep track of the source of the data and how it has been treated. One possible solution is to connect BI tools directly to operating systems. This can be a reasonable solution in small and simple environment where the amount of data sources is very limited. Problems are faced when there occurs need for compiling data from various sources and there is no storage for historical data. Virtual DW is also one possible solution for making data structured and applicable. This can be seen as a tempting solution especially from a business point of view. There is no need to invest in DW and the data can be compiled and structured on the way to reports. It has remarkable problems in compiling and structuring data. This process requires complicated deduction logic and parts of the data cannot be examined on its own. Also, it lacks historical dimension of reports and therefore it is impossible to produce time series or trend analysis. (Hovi, et al., 2009, pp. 6-9) DW solutions provides an answer for this need to store, process and compile information from various sources. Data warehousing delivered a possibility to build BI systems. Without the underlying data warehouse BI was just a theory. (Inmon, 2005, p. 402)

DW as a term was introduced for the first time by (Devlin & Paul, 1988) IBM Systems Journal. It was aiming to provide a working definition for discussion about data storage and processing. William H. Inmon has been considered as one pioneering researcher and contributor in the field of DW according to (Hovi, et al., 2009, p. 11). Indeed, the first edition of his first book “Building the Data Warehouse” was published in the 1990 and it aimed to distinguish DW as an architecture, not a technology. The subject was controversial and not respected notion at the beginning in academic and business circles. Today DW is not just a theory but a fact of life according to (Inmon, 2005, pp. XX-XXI). Later on Inmon has gained reputability with tens of thousands of citations to his publications in academic research papers.

Even though BI as a term was defined several decades ago it became more popular in the business and IT communities in the late 1990's. Information and data sources were usually internal systems, customer information and financial data. Analytical techniques built on top of these relational databases are mainly based on statistical methods. These kind of BI solutions are often referred as BI 1.0 (Chen, et al., 2012, p. 1166; Olszak, 2016, p. 107) Next step in the BI technique development came from advanced DW techniques and data mining techniques. These new web technologies allowed organisations to process and analyse unstructured data from the web more efficiently. Commonly, these data sources are organized, combined and visualized to reveal behavioural customer patterns. These advancements are referred to as BI 2.0. (Chen, et al., 2012, p. 1167; Olszak, 2016, p. 107) The latest advancements, now dubbed BI 3.0, present a new era in the evolution of BI. This development incorporates innovations including utilisation of exponentially growing amount of data, mobile device related analytics, sensor-based devices connected to the Internet, location intelligence and cloud based BI services on demand. It has been seen that these additions brought in by BI 3.0 do not exclude functions from earlier BI innovations. To summarize, the main philosophy is to increase the added value of BI tools. (Olszak, 2016, p. 107; Chen, et al., 2012, p. 1168)

During the last two decades BI&A accelerated in tremendous growth both in academical and business fields. It is a potential business enabler across industries, and it is one of the fastest growing fields in business. (Baryak, 2015, pp. 230-231; Davenport & Dyché, 2013, p. 1) At the beginning of 21st century key research themes regarding BI, BA and BD came up only in a few studies while one can find over 2500 studies in just one database with these key research themes in 2014 (Baryak, 2015, p. 234). This trend in number of studies regarding these mentioned key research themes can be seen also in many other scientific databases. For example, (Marjanovic & Dinter, 2017) confirmed a similar booming trend in the Hawaii International Conferences on Systems Sciences (HICSS) database publications with key research themes in above mentioned BI, BA and BD field. Notably (Chen, et al., 2012, p. 1179) named HICSS as the premier conference for BI&A research. Perhaps surprisingly, BI and its applications are not limited only to a specific field of science. In fact, BI is applicable in many different fields like Mathematics, Geography, Social Sciences or

medicine. Top three fields for these key research themes were not a big surprise as one would imagine the possible research settings. These three fields in the following order were: Computer Science, Business & Management and Engineering. It seems that studies related to this field are not just a trending topic rather a driver for changing the way of decision-making and predicting future outcomes. (Baryak, 2015, pp. 237-239) Also, interest towards BI has continued its growth in the business world too. Firms have been reporting more competitive value gains from the use of BI&A in their business. In addition to that, there is evidence that businesses continue to make IT related investments. It seems that business leaders have identified the possible benefits and competitive gains that it can offer. (Božič & Vlado, 2019, p. 93; Nordic Institute of Business & Society NIBS, 2020, p. 16)

2.2. BA, BI and BD

BA can be defined as *"a broad category of applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users make better decisions"*. Terms BI, BA and BD are used as synonyms or referring to similar subjects in academic and commercial context. For an example IT community is more familiar with using term BI whereas business community prefers using term BA. (Baryak, 2015, p. 231) For an example (Davenport, 2006) used BA term to emphasise key analytical components in BI. It seems that the term BI does not have an unambiguous explanation and the use of the term is case sensitive.

In many cases BI has been defined as a tool or managerial concept that is used for managing and leveraging business information to support operative and strategic decision-making. This process produces up-to-date knowledge and intelligence about the underlying business. (Ghoshal & Kim, 1986; Gilad & Gilad, 1986) In these two articles the explanation of this term is presented as two dimensional:

1) **Processed information and knowledge** that is based on data. Information is acquired from the surrounding environment and inside the business
2) **The process itself** which transforms data into insights, suggestions and recommendations for the decision-makers

Remarkable parts of BI processes are connected to the underlying data which is the backbone for the whole process. Modern BI solutions are sophisticated IT applications which processes and analyses the data. Handling and storing this data is considered as an important part of BI. (Moss & Atre, 2003) Hackney (2000, pp. 39-42) gives a similar definition for term BI in his article. The main idea in BI is identifying information needs and then processing gathered information and data into valuable and utilizable form to gain managerial intelligence and knowledge. Remarkable thing in his research is that he solely focuses on company's internal information.

Another definition for BI is presented by (Hovi, et al., 2009) where the explanation of term BI still has two dimensions in it. This explanation approaches the term form data source point of view, see Figure 5 below. These are often referred as quantitative outlook and qualitative outlook. Quantitative outlook is referring to BI applications which use data from internal sources. For this kind of information sources is typical that the data is in structured and easily storable form. The early BI applications usually relied on these internal information sources. Qualitative outlook emphasises external sources of information in BI applications. This kind of data sources are usually unstructured and require more processing before utilization. Qualitative BI applications are named with terms such as Market Intelligence (MI) and Competitive Intelligence (CI) which describes more accurately how these processes are utilized. (Hovi, et al., 2009, p. 78) These two above stated interpretations for term BI are not scientific and accurate definitions and it does not include all academic research done around this topic. This can be seen as a distinction between the applications of BI and in which kind of business-related questions it is designed to answer.

Information form			
Structured	Market data (stock prices, etc.)	Business Intelligence (Sales-, production-, personnel-, financial information)	
Unstructured	Market Intelligence (News related to the operating market, competitor information and news)	Enterprise Content Management (Document management, electrical invoicing system)	
	Outside	Inside	**Information origin**

Figure 5 Two interpretations of term BI after (Hovi, et al., 2009, p. 79)

As one can see that BI is built highly reliable on the underlying data. This leads the discussion closer to the field of BD and therefore the definition of term BD should be discussed. When talking about BD in this context it refers to complex and big DWs that companies have. Companies have built these DW solutions as an answer for the increasingly growing amounts of data. (Baryak, 2015, p. 231) And talking about big databases, the size of these large databases is in the peta-scale (one petabyte equals thousand terabytes which is thousand gigabytes) and they are growing all the time. These DWs holds structured data and non-structured data such as plain text files. (Hovi, et al., 2009, p. XIII; Inmon, 2005, p. XXV) Data management and warehousing can be considered as the foundation for BI and therefore there is a need for modern data processing and analytical technologies. (Chen, et al., 2012, p. 1166)

Probably the first thing that come into one's mind when talking about term BD is large masses of data. Besides that, it has more characteristics than just the amount of data. Laney (Laney, 2001) suggested that Volume, Variety and Velocity are the three dimensions that describe the challenges in data management. This proposal has emerged as a common framework for describing big data. It is widely known and called as The Three V's. (Gandomi & Haider, 2015, p. 138; Chen, et al., 2012, p. 1182) Later, in the year 2012 IBM presented additional dimension to describe BD. Fourth V comes from Veracity and it has been widely adopted as a part of describing BD. (Schroeck, et al., 2012) These four V's of BD can be presented as follows in Figure 6.

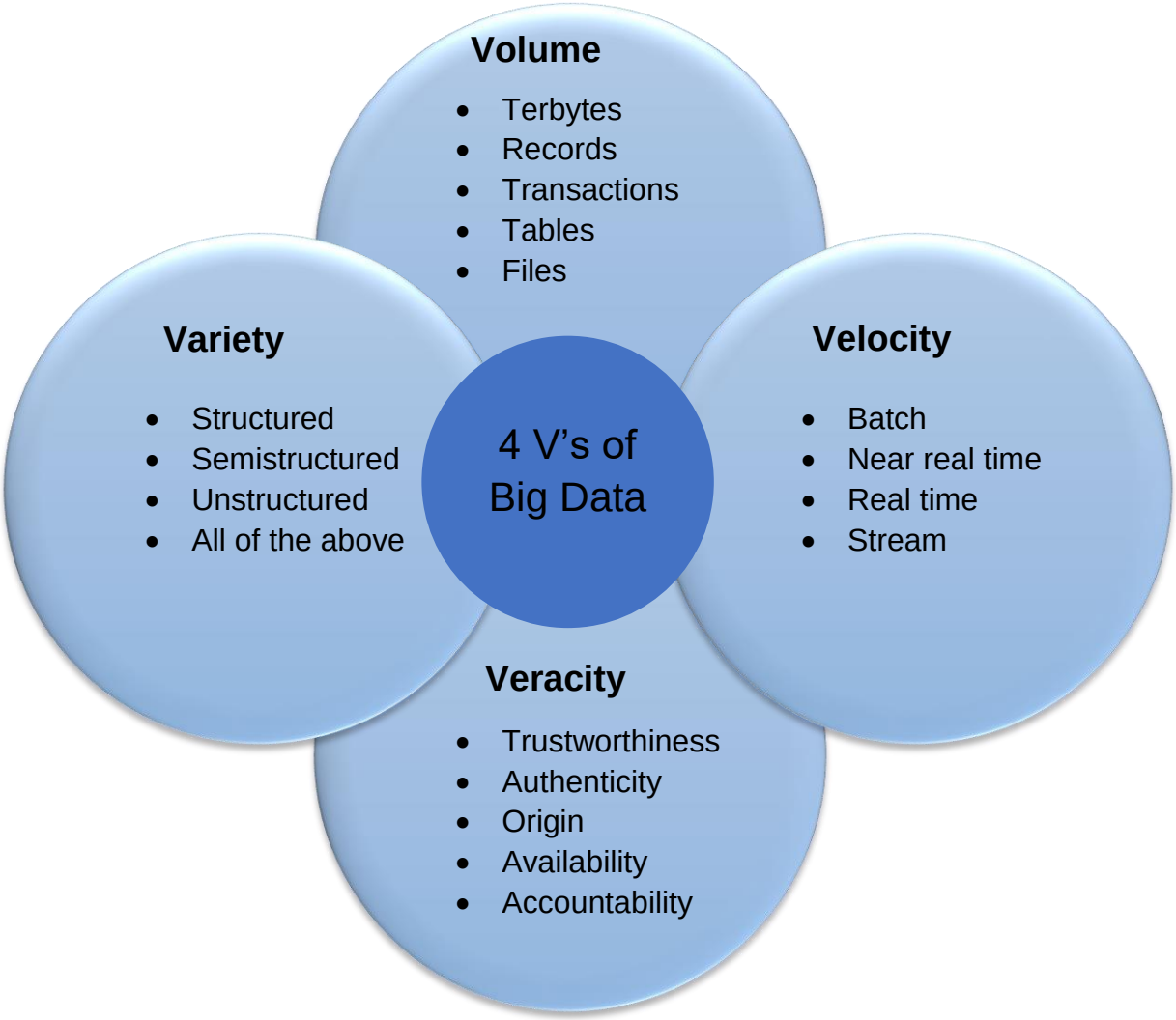

Figure 6 4V's of Big Data after Laney (2001) and Schroeck, et al. (2012)

Volume refers to the magnitude of data and the size of BD is usually reported in terabyte or petabyte class. (Gandomi & Haider, 2015, p. 138) The amount of available information in organisations has increased tremendously. This has made even more challenging to identify the relevant information that should be considered in decision-making process when data volumes become large and complex. (Baryak, 2015, pp. 232-233) Variety refers to the structural heterogeneity in data. BD combines various types of structured, semi-structured and unstructured data. Structured data refers to tabular data found in spreadsheets or relational databases. Text, images, audio and videos are examples of unstructured data. One example of unstructured data is XML

files in internet that contains tags defined by user to make it machine-readable. (Gandomi & Haider, 2015, p. 138) Velocity of data comes from the rate at which data is generated and how fast it should be analysed and act upon it. Most data scientist struggles with the data velocity to make fast and still accurate decisions based on data. (Baryak, 2015, p. 233) The rise of data transmitting and processing digital devices such as smartphones and sensors is driving a growing need for real-time analytics and faster response from the information produced. (Gandomi & Haider, 2015, p. 138) Veracity refers to the level of reliability on certain types of data. In order to make accurate predictions and create holistic reporting with BI the quality of data must be assessed. Uncertainty of the surrounding and some data sources need to be acknowledged by the decision-makers. Examples of this kind of data can be economic data, weather data or customer's actual future buying decisions. (Schroeck, et al., 2012, p. 5) Discussion about data quality and its methods is a whole scientific field on its own. In this research, all aspects related to underlying data quality are ruled out.

It has been argued that the ability of using BD in decision-making process leads to competitive advantage against rivals. Some of the earlier studies have found that it has enabled companies to achieve outstanding performance comparing to their peers. However, it seems that there is lack of analytical techniques for utilizing these data masses and gaining competitive advantage. (Tan, et al., 2015, p. 223) BI has been considered as an umbrella term that contains architectures, databases, tools, applications and methodologies. All these above-mentioned parts aim to analyse and produce insights from the data to support decision-making process. (Moro, et al., 2015, p. 1314) Many of the most cited authors in the academia seems to use BI as an umbrella term of Decision Support Systems (DSS). And there is evidence that BI is used to describe these same systems in business also. (Arnott, et al., 2017, p. 58; Olszak, 2016, p. 107) This literature review conducted by Olszak unveils the truth that BI has not been defined unambiguously and it can refer to many different applications and methods. Different aspects that have been considered to be a part of BI definition are presented below in Figure 7.

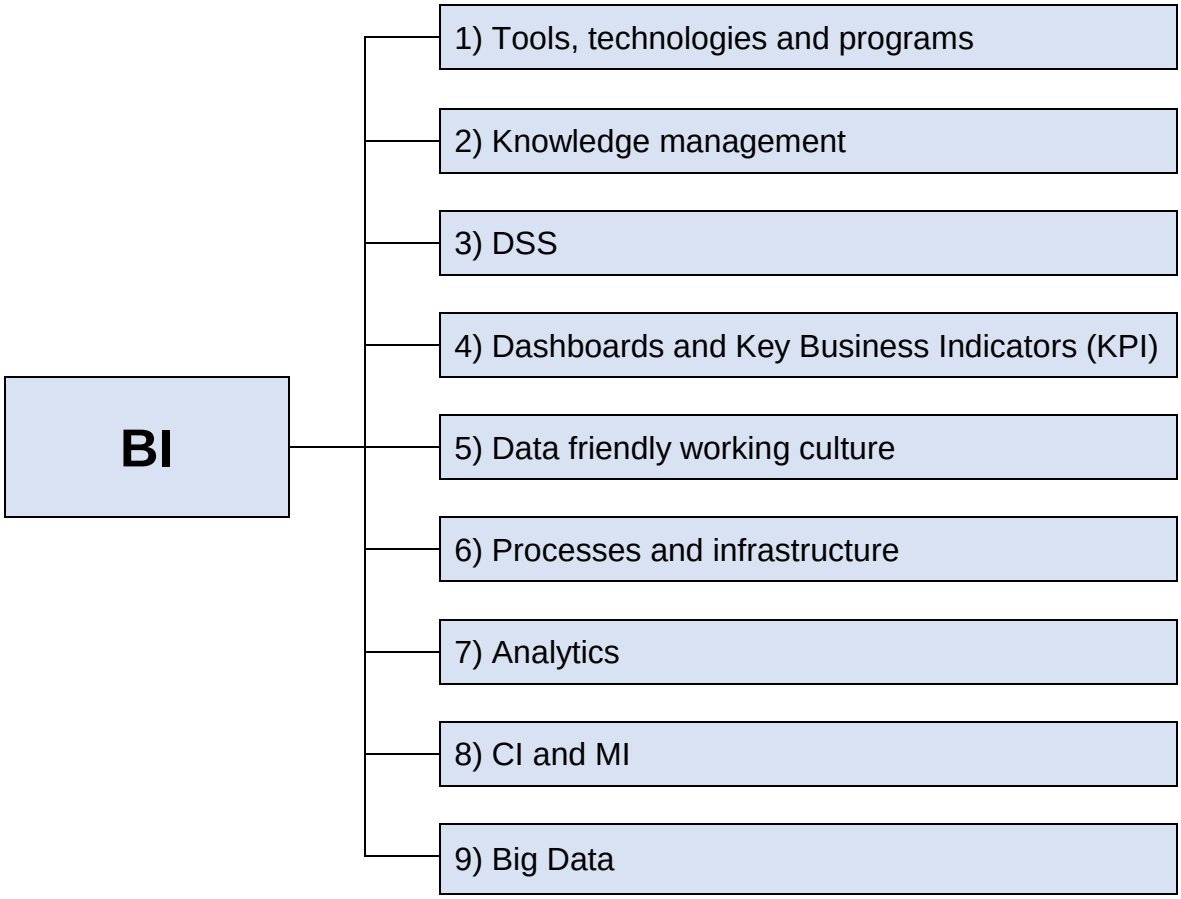

Figure 7 What lies under term BI after Olszak (2016)

Although it seems that there are ambiguous definitions for the term BI. Besides the differences between these different explanations, all definitions do interrelate together in a way or other. In this thesis BI is understood as **a systematic process of intelligence for obtaining data to serve the information needs of interactors and decision-makers to achieve insights and competitive advantage.** This process can include gathering, manipulating and combining data from sources inside or outside of the company. This data can then be stored and analysed for further use. By automating this process, it is possible to get more accurate and timely information that is essential for business activities and helps in the decision-making process. Even though BI is considered in this study as above stated the earlier mentioned components of intelligence in this paragraph can be seen as a different viewpoint of BI. These different viewpoints are left out of the scope of this study and more emphasis on the aspects of BI are given that are related to the definition presented in this thesis.

2.3. Benefits

Thierauf (2001, pp. 3-4) emphasizes that BI can deliver valuable information for a company to utilize in business. But he also remarks that information obtained with BI process is just data until it has been utilized in analytical purposes, decision making or gaining insight. Benefits of the BI has been studied in the academia multiple times from different viewpoints. For an example (Scholz, et al., 2010, p. 11) found in their study that there was three commonly found BI benefit factors. First factor was improvement in data support which included overall increase in quality and competence in data analytics around reporting. The second factor was related to decision support. Companies experienced improvement in decision making process when it was backed with data and identifying possible risks and changes in operating environment was seen to be easier. Third factor was savings in operating costs from more efficient resource allocation inside the company. Similar operational and tactical improvements have been found in other studies too (Elbashir, et al., 2008, p. 149).

Davenport and Harris (2007) shares this same view of BI being the enabler to transform data into valuable insights in their book. Many companies use analytics to gain edge and operate business with maximum efficiency and effectiveness (Božič & Vlado, 2019, p. 101). The usefulness and power of BI was illustrated with multiple business application examples across industries. Just to name few examples presented in the book written by Davenport and Harris were such as Netflix, Google, Amazon, insurance and credit card companies. These companies have used analytics to understand their customers better and trying to recognise and retain the most profitable customers. To describe the level of competitive advantage and degree of intelligence they divided BI process in data access and reporting, and analytics. See Figure 8 below. Each of these steps provides information about an organisation’s business activities. Analytics built on top can answer to more higher value questions and provide more proactive solutions.

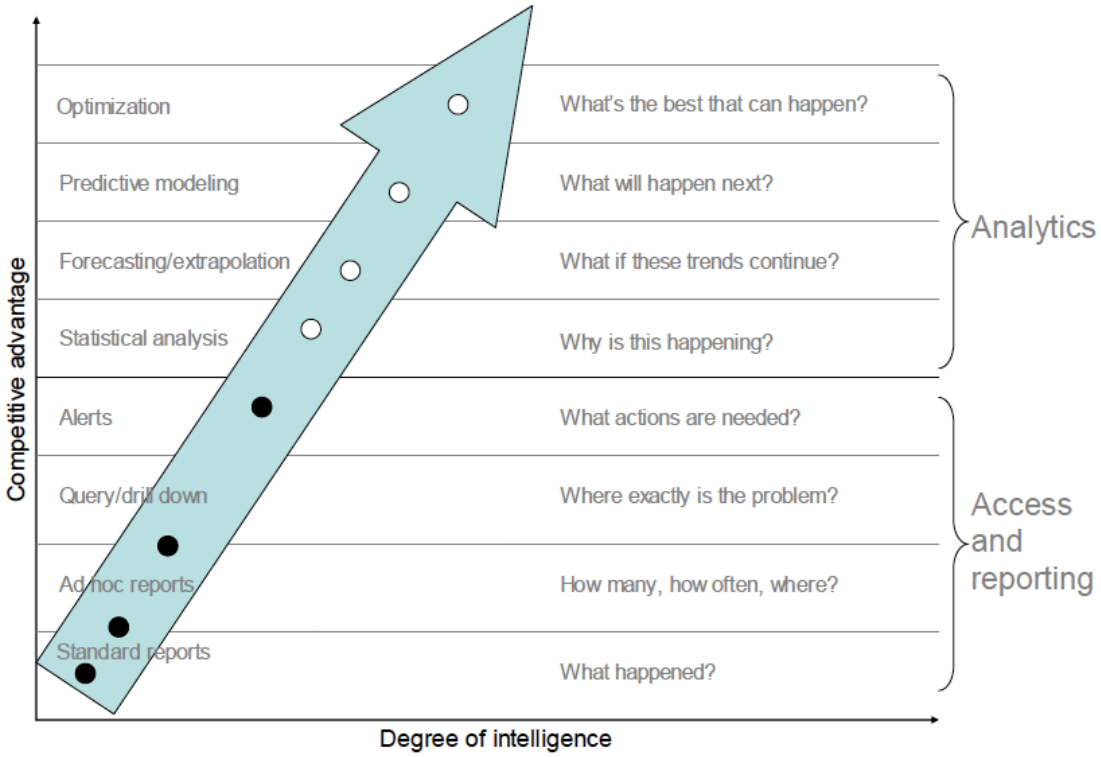

Figure 8 Business Intelligence and Analytics *(Davenport & Harris, 2007)*

Due the intangible nature of BI activities and its outcomes it is a hard task to measure financial benefits from it. Lönnqvist and Pirttimäki (2006) have summarized two types of BI measurement methods distinction in a similar manner as earlier discussed two-dimensional definition for term BI by (Ghoshal & Kim, 1986; Gilad & Gilad, 1986) at the beginning of Section 2.2. The first and most common reason for measuring BI comes from urge to prove that it is worth the investment. In many cases is needed to demonstrate why BI is needed to process data to insights and actions. The second reason for measuring BI process is to manage the process itself. This is done to ensure the efficiency of BI process and verify that the process fulfils the user's needs. Even a sophisticated process or big investment can be pointless if the gathered information is not accurate or does not match the information needs. (Lönnqvist & Pirttimäki, 2006, p. 33) These two main measurements of proving BI benefits have been summarized below in Table 1.

Purpose for measurement	Main users of measurement information	Expected benefits
Valuation of the worth of BI	Companies (i.e., executives) applying BI BI service providers BI professionals Researchers	Ability to prove that BI services are worth the effort and demonstrate the actual effects of BI. Increased credibility of BI as a managerial tool. Improved rigor in BI research.
Management of a BI process	BI service providers BI professionals	Continuous improvement of BI products and services

Table 1 The types of BI measurement *(Lönnqvist & Pirttimäki, 2006, p. 33)*

Lönnqvist and Pirttimäki remarks in the article that the measures for managing the BI process have not been discussed as widely as measuring value and effects of BI (2006, p. 35). The steps of describing the level of competitive advantage and degree of intelligence presented by Davenport and Harris earlier in this paragraph can be seen to belong more under the management of a BI process. Questions presented by different steps can make it easier to justify the BI process and direct it to answer the information needs of a company more accurately. One way is to evaluate company's BI readiness in sense of a balanced scorecard system from various aspects. For an example technological readiness, strategic alignment, data analytics culture and striving for improvement inside the company are measured to manage BI process (Williams & Williams, 2004, p. 5) This method seems to be quite straight forward and requires self-reflection abilities inside the organisation. One can look closer its own actions and abilities to discover more innovative and efficient ways gaining benefits from BI processes.

While defining the level of sophistication of BI process many studies have been seeking an answer to quantify the value that BI has generated to the company. When determining the valuation of BI one must consider the value for whom? In their article Lönnqvist and Pirttimäki gives out two main questions when assessing the value of BI: *"How much does it cost to apply BI?"* and *"What are the benefits of applying BI?"*. The costs of applying BI is much easier and can be determined as like any other investment costs. But measuring the benefits of BI are not as simple as measuring the costs. (2006, p. 34) When studies have found implications about improvement in decision-

making and operational performance one could expect that there would be evidence about increase in financial performance too. And this seems to be the case. There has been conducted a study which created a measure of the use of data driven decision making on productivity, financial performance and market value. They examined and tested this hypothesis on 179 publicly traded firms in the US and confirmed a 5-6% increase in the company output and productivity beyond what could be explained with traditional inputs and IT usage. (Brynjolfsson, et al., 2011, p. 541) And still, this seems to be the case in wider perspective too. (LaValle, et al., 2011, p. 22) conducted a survey containing nearly 3000 executives, managers and analysts across over more than 30 industries and 100 countries. Their study found a clear connection between performance and the utilisation of BI to gain edge. Notable finding was that *"organizations that strongly agreed that the use of business information and analytics differentiates them within their industry were twice as likely to be top performers as lower performers"*. It can be seen that there is clear evidence of value generated by BI to companies across industries. The level of intelligence and valuation is possible to measure but it is not a simple and straightforward task in many cases.

It is evident that there are numerous benefits from BI as one can see from the studies presented above. These benefits can be organisational in a sense that the company's internal processes are improved. It is also evident that use of BI has a positive impact on business performance. One common thing with these studies is that all they underline that BI can transform data into valuable insights that can be utilized in decision-making process. And coming to discussion about measuring the benefits of BI there are many different ways to do it. These different methods can be divided in two categories as presented in Table 1 by (Lönnqvist & Pirttimäki, 2006). These categories are determining the value of BI and managing the BI process itself. The aim of this study is not to value the BI process, rather looking closely the process itself. In the following section this process itself is examined more in detail.

2.4. Process

The objective of a BI process is to transform data and information into knowledge and insights. The process itself can either be for a one-time ad-hoc purpose or more continuous and systematic in nature. (Gilad & Gilad, 1985, p. 66) With ad-hoc analytics

a company can analyse and gain information of one specific situation. One possible example would be a scenario analysis on the effects of certain factors in revenue to support managers in decision-making. These factors include e.g., product price, sales volume, sales margin or production costs. Systematic BI process enables a company to gather information continuously. This information can be used to steer the business and make decisions based on changes originating from in and outside of the company.

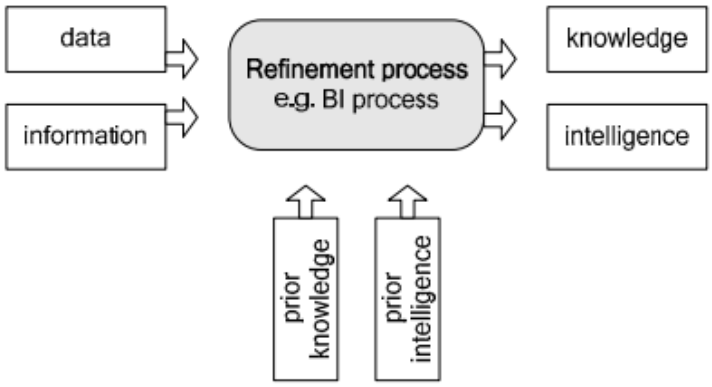

Figure 9 Transforming data and information into knowledge and intelligence after (Hannula & Pirttimäki, 2005, p. 36)

(Hannula & Pirttimäki, 2005) illustrated this process as presented in Figure 9. Data and information are the main inputs for the BI process in producing knowledge and intelligence. Prior knowledge and intelligence are a significant part of the process. An emphasis is given on the human insights and experiences because learning is one critical part of BI process improvement. Therefore, continuous BI process can be seen as a cyclic system repeating itself. This leaves room for continuous improvement and adaption. One possible way is to present BI process model as a cycle with different phases like in the following figure.

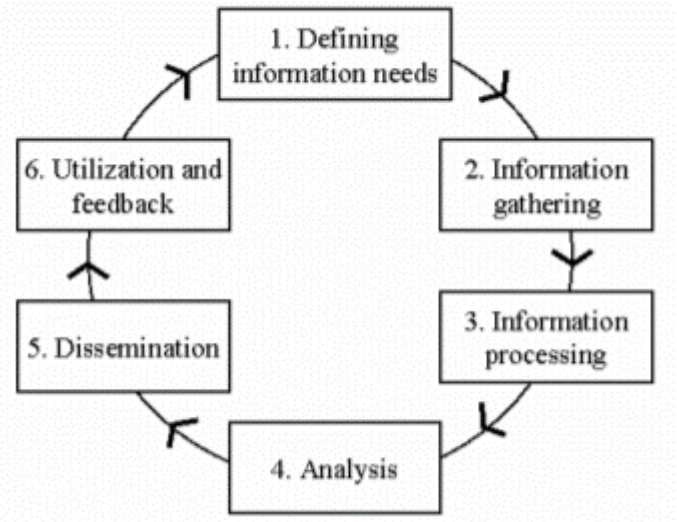

Figure 10 Generic Business Intelligence process model after (Vuori, 2006, p. 2)

The number of phases in the cycle can vary a bit in the literature. One possible way of presenting this is to generalize the process to six clearly recognizable parts. The process starts with identification of information need which is always in common with different process models of BI. This phase is followed by steps to gather, store and process the information. After these steps, an analysis with the data is performed. Then the findings are distributed and communicated amongst the user group. The last step includes the utilization and assessment of the results. It also leaves room for the feedback for continuous improvement and fine tuning of the process. (Vuori, 2006, p. 2) Similarities to this generic BI process model are clearly visible in other academic papers too. For an example (Choo, 2002) presented an illustration of BI process in his book which concentrated on the information management aspects.

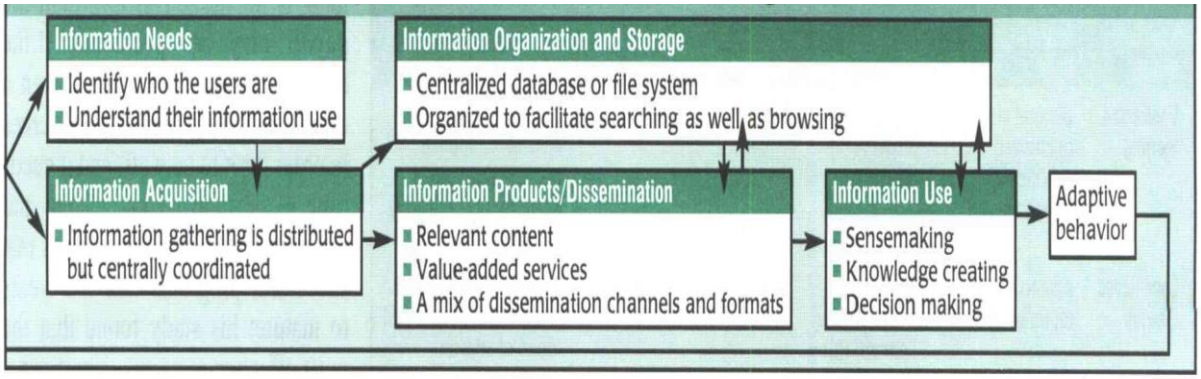

Figure 11 Process model of information management after in environmental scanning after (Choo, 2002)

The BI process definition presented in Figure 11 builds around information management. There are six stages identified in this process which are similar to the steps presented in model by (Vuori, 2006) in Figure 10. Both models begin with defining information needs and proceed to information collection. Information storage comes as the third stage where the information is in accessible, shareable and usable form. These make up the following stages of this process. Notable is how arrows present interrelation between the boxes. This enables utilization of prior knowledge and intelligence in the process which is similarly presented in the article by (Hannula & Pirttimäki, 2005), see Figure 9. Last stage in this model incorporates the adaption to improve the process while repeating this cycle. This conceptualizes entire information value chain from data to information, knowledge and decision-making. (Choo, 2002, pp. 24-25)

When going through newer literature about BI process models it seems that similar elements are recognized to be part of such a model. Same themes and emphasis are surfacing from study to study. The number of steps may vary but the main idea remains the same. For an example (Pellissier & Nezhelele, 2013) proposed a universal cyclic process model for CI similar to the generic model presented by (Vuori, 2006) in Figure 10. (Hellsten & Myllärniemi, 2019) revisited BI process models in their article and found the process model of information management presented by (Choo, 2002) to be still relevant in this field, see Figure 11. They mentioned that the general cyclic BI process model was also still relevant. As a conclusion they phrased that the BI process itself is

complex with many moving parts. This complexity comes from the changing environment, unique setting of the business and increasingly demanding sophistication level of BI tools. (Hellsten & Myllärniemi, 2019, p. 347)

When these different steps from various BI process models are taken into account following general BI process model pinpointing steps found the most relevant can be proposed. It is worth underlining that the model is constructed to be cyclic due its nature. This leaves room for self-reflection and evaluation after every iteration. Self-reflection and evaluation can be seen similar to the organizational lean practices. It allows to learn from own mistakes, try new practices and adapt to changes along the way. This general BI process model can be presented in five steps as follows.

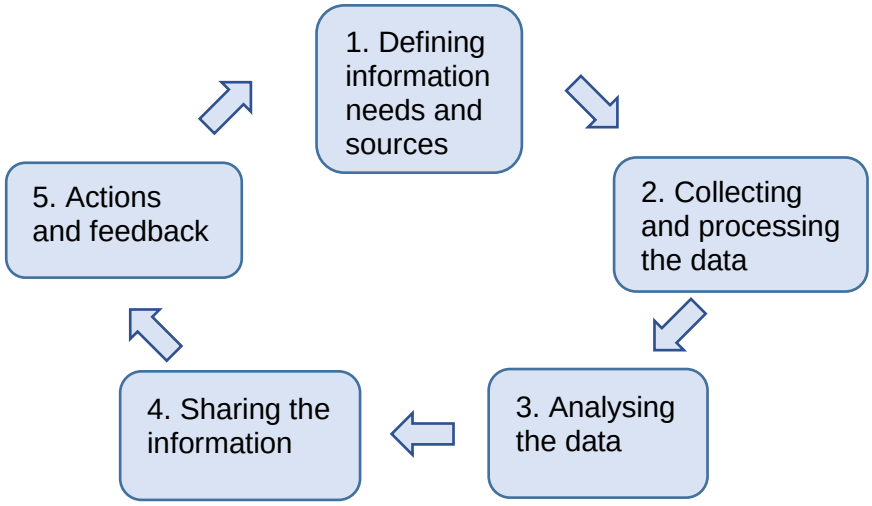

Figure 12 General BI process model

The first step of this model plays very crucial role in the successfulness of the whole BI process. In order to transform data into insight and gain the benefits discussed in the Section 2.3. One must start with defining intelligence needs, data sources and priorities for the system. This step has been described as a buffer preventing managers to swamp in overload of information in the decision-making process. Therefore, it is necessary to define the information needs and relevant data sources. (Gilad & Gilad, 1985, p. 66; Hellsten & Myllärniemi, 2019, p. 344) Acknowledging the fact that businesses have limited resources it is ultimately important to distinguish the actual

information needs from irrelevant “nice to know” type of information. In addition to the essential information that the business already has it must identify possible information needs. These needs can be split into three categories: *1) Information that is wanted but not actually needed.* Obtaining irrelevant information is a waste of resources. These misbeliefs usually originate from poor understanding of the actual underlying problem. *2) Identified need of a missing information.* Resources and effort used on obtaining needed information is essential work of a BI unit in a company. This information can be used in value creation by turning data into insights. *3) Information that has not been asked or known to ask for.* Relevant information may have been left out from the BI process and it may change the outcomes from the analysis. (Vuori, 2006, pp. 4-5) It is important to find the right questions to ask from the data. This can be hard and iterative process achieved by trial and error.

Collecting and processing the data is the second step in the model. This step is strongly driven by the identified information needs and the sources of this information. Data collection can be considered as a complex function because the data collected from in and outside of organization can be in various formats and tricky to acquire. (Lönnqvist & Pirttimäki, 2006, p. 37) Collected data must be stored in a way that it is in readable and usable format for the BI applications. This requirement creates an opening for Extract Transform and Load (ETL) process as illustrated in Figure 4 at the beginning of Section 2.1. ETL pipeline acts as a middleman in the process of collecting and processing the data between data sources and DW.

Third step of the model is data analysis phase. In this step the collected and processed data and information is being transformed into useful knowledge and intelligence. This refinement process presented by (Hannula & Pirttimäki, 2005) was discussed earlier in this paragraph. Processed information aims to answer the needs of the decision-makers and provide insights. BI tools assisted with analytical and statistical methods come in handy at this point. This part can be described as applying mixture of number-crunching, technology and analytics on business problems. By doing this companies can churn value from their data. (Hellsten & Myllärniemi, 2019, p. 344; Davenport, 2006, p. 103)

Having the right information and insights does not make any company wiser or more competitive if the outcomes are not communicated. The fourth step of the model is about dissemination of the information and insights. Results from the analysis step needs to be delivered to the right recipient in a timely manner with relevant methods (Hellsten & Myllärniemi, 2019, p. 344). These results are the finished product from the process itself and is actionable intelligence. It is important to communicate these results in a way that is easily understood. Usually in the form of a report, dashboard, intranet solution, email or meeting. (Pellissier & Nezhelele, 2013, p. 4) It is advised to consider the personal routines and preferences when planning the information delivery. Wider information sharing that is understood more in detail can lead to positive outcomes in the organization. Organizational learning curve steepens, retrieval of information to answer business problems comes more frequent and new information needs can be discovered. (Choo, 2002, p. 44) These remarks leads towards the final step in the model.

Fifth step includes the actions taken by decision-makers and collecting feedback from the process. Information acquired in the process are turned into action in problem solving purposes and decision-making situations. There hardly exists any source in the literature that would not emphasize the importance of this step. Utilizing the acquired information and knowledge can be used to adjust the process and the BI process cycle starts over again. (Hellsten & Myllärniemi, 2019, p. 344) This prior knowledge and information have been recognized by (Hannula & Pirttimäki, 2005) to be a crucial input for the BI process as shown in Figure 9. Constructive feedback from the users of BI enables a continuous improvement atmosphere (Pellissier & Nezhelele, 2013, p. 5). Learning from the process and experimenting with new features is familiar from the lean practices. This self-reflection leaves space for improvement and organizational growth.

This general BI process model presented in this study is applicable to various use cases in organizations across different industries. It can be followed to identify the steps in a company BI process and unveil possible challenges in it. The next section of this study introduces the common infrastructure behind this process.

2.5. Infrastructure

Infrastructure built for the BI process allows to collect the information, manipulate the data and store it for utilisation purposes. These needs were identified in the previous section where the general BI process model with five steps was presented. Automating these data collection and preparation steps can be considered as an important action to make BI more effective. Advancing closer towards real-time BI is a topic to tackle in the near future. These features have surfaced in the studies for a while but are emphasized more in the recent studies. (Hellsten & Myllärniemi, 2019, p. 345) Automation of these steps is possible with a suitable infrastructure. In this paragraph technology behind BI processes is discussed briefly without going into finest technical details.

Data from numerous systems and sources are collected, stored and treated systematically in BI process for the needs of decision-makers. It is also important to assure that no critical information is lost, information gaps are identified and covered, duplication of data sources avoided, and all relevant information gathered to support strategic planning. (Choo, 2002, p. 91) In many organisations BI has a gatekeeper role between the overload of information and decision-makers by filtering out all the non-relevant information from decision-making process. (Vuori, 2006, p. 2) Architecture for such systems have been typically described with three layers as the following figure illustrates.

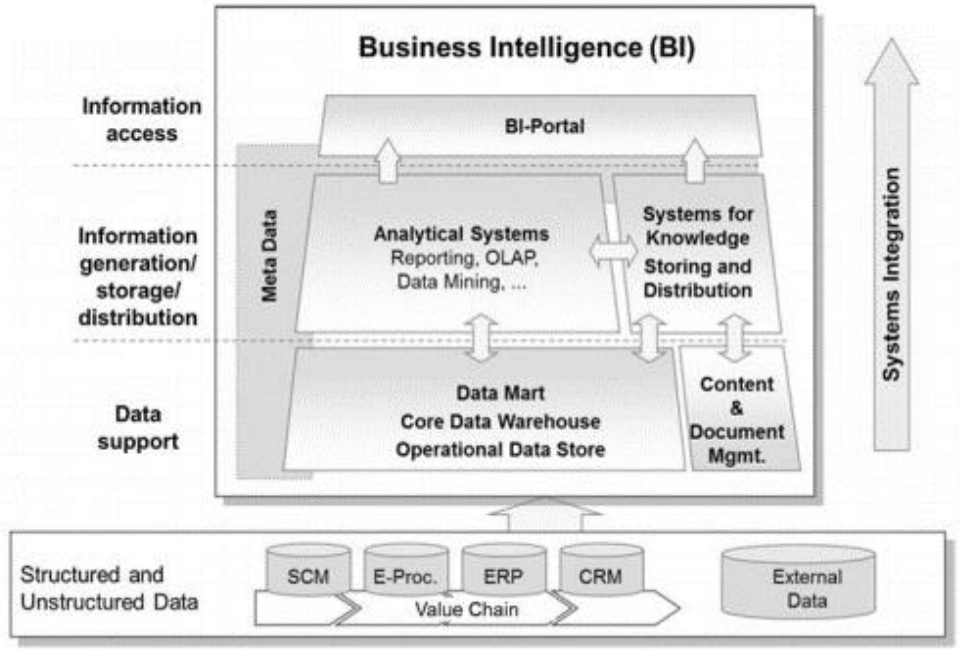

Figure 13 Business Intelligence Architecture (Rausch, et al., 2013, p. 5)

Bottom line for the architecture and BI systems are the data sources from operational systems like Supply Chain Management (SCM), E-Procurement Systems, Enterprise Resources Planning (ERP) Systems, Customer Relationship Management (CRM) and other external sources. ETL pipelines are used to transfer data from these sources to the Data support level. At this level data is stored in DW, treated and harmonised for further utilisation in BI process. (Rausch, et al., 2013, p. 5) This bottom layer of BI architecture enables the data collecting and processing activities described in the second step of the proposed general BI process model (see Figure 12 in Section 2.4.). Bottom layer can be seen as a foundation for a working BI solution.

Next layer includes the ability to analyse the collected data, store the outcomes and support the distribution of information and knowledge. Analytical features of this layer include such as Online analytical processing (OLAP), data mining, (interactive) reporting and ad-hoc analysis. (Rausch, et al., 2013, p. 5) This layer has similar attributes as described in the third step of the proposed general BI process model about analysing the data.

Third, the top layer of this architecture model includes the information access related activities. It offers access to all relevant BI functions in unified environment with defined user roles and rights. Usually, this layer is implemented with graphical user interface suitable for communicating and exploring the information and insights. (Rausch, et al., 2013, p. 6) One can clearly notice similarities to the fourth, information sharing step in the proposed general BI process model. Notable thing is that modern BI solution architecture normally consists of a variety of different data sources, legacy systems, diverse ETL pipelines, domain-specific data granularities and adept analytical tools. BI process models are used to tackle the challenges in the designing phase of decision-making supporting activities. (Rausch, et al., 2013, p. 6)

As mentioned earlier, this study will not go deeply into technical details of BI architecture. Nevertheless, it is important to understand the basics of BI architecture when aiming to improving and automation of the BI process (Rausch, et al., 2013, p. 9). There are strong arguments in favour for suitable architecture when building BI process and automation. It is more forward-looking tactic than just trying to parse something up from bits and pieces. On top of a good infrastructure is possible to build BI&A applications more rapidly and easily. Well maintained architecture allows company to try different analytical approaches and explore the data. It leads to more versatile BI solutions. This increases productivity and data can be refined more and more. Business can gain competitive advantage from data and help in decision-making process. If data management is done properly the advantages are continuous and scalable. These advancements allow company to use DDD. (Božič & Vlado, 2019, pp. 100-101; Rausch, et al., 2013, p. 24) Now when the theoretical framework and background for this study is presented it is time to move on to the literature review part of this study.

3. Literature review

Now, when an overview about the theoretical background of BI and its closely related concepts has been built it is time to move on to the literature review part of the study. Based on the research questions proposed in the first chapter of this thesis a literature review needs to be carried out. This literature review fits in the research focus presented in Chapter 1.2. and helps to answer the second sub research question. The second sub question was linked to state-of-the-art academic literature about BI use cases and faced problems in the implementation. Looking at what has been written about BI automation use cases in the academia helps with comparing earlier research to the observations and findings later in the case study section. Based on the findings from previously conducted studies the focus can be targeted more closely on relevant aspects of the BI process automation in the case study. This will also help in forming proposals to improve the observed company implementation practices. In fact, this is closely related to the third sub question and gives a good support for answering to it.

Literature and articles used in this study was found from various scientific databases found in LUT Finna: EBSCO – business source complete, Elsevier Science Direct, IEEEXplore, Scopus and Google Scholar. Some of the material was also found from the references in other articles and books related to BI. The research results were narrowed down to literature / articles that were accessible and related to BI or its implementation.

3.1. Use case examples of BI

It is found ideal to summarize and present the found articles in one table. These articles were split into two categories: Newer BD / IoT solution based BI cases and more traditional BI cases. Splitting the found articles in two categories by the technology comes from the latest innovations in BI. BD and IoT solutions are one of the newest advancements in the field of BI. Also, the observed case in this study falls under the more traditional BI solution category. After reviewing the articles a table comparison is made to summarize the use case types and challenges faced in the implementation.

BI solution type	Authors	Use case types	Challenges faced
BD and IoT based solutions	(Ahmed, et al., 2015), (Stockinger, et al., 2015), (Moro, et al., 2015), (Ramalingam & Venkatesan, 2019), (Gyulai, et al., 2019)	Analyse transaction data in retail industry, IoT enhanced factories, Outsourced infrastructure management, Credit risk evaluation, bankruptcy and fraud prediction in banking industry, IoT applications in banking to provide closer to real-time analytics, IoT applications to make even more detailed decision support systems.	Handling large data masses, Need for higher processing and analytical performance, Network dependency issues when providing closer to real-time analytics, Not relying enough on data in decision-making process.
DW, OLAP and other more traditional solutions	(Gangadharan & Swami, 2004), (Bhatti, 2013), (Brichini, et al., 2015), (Ally & Khan, 2016), (Banda & Ngassam, 2017)	Manufacturing process planning and optimisation, Compiled reporting and monitoring from separately acquired business units, Using BI to manage existing BI systems, Data and information-based decision-making tool for health sector, Trend analysis tools to support decision-making process.	Creating an infrastructure capable of handling legacy systems and multiple data sources, Ensuring the data quality when combining results from various sources, Difficulties to identify, define and validate right indicators and measures, Restricted culture of sharing information inside the organisation, Issues with data latency.

Table 2 Summary of observed articles in literature review

By looking at the summary it can see that BI has multiple applications across industries. These findings support the remarks made in Section 2.1. that BI is not only limited to a specific field (Baryak, 2015, p. 238). One can recognize familiar aspects from these articles that surfaced in the theoretical framework section of this study. Common use case types include the aspects of transforming data into actionable information and

insights to support decision-making in organisations. All these processes have steps with data collecting, processing, storing, analysing data and disseminating the outcomes as presented in the generic BI process model presented in Section 2.4. It seems that different types of processes that produce data frequently can be automated and applied into BI systems. Next, in the following paragraph BI use cases in banking sector are being inspected more closely.

3.2. Banking and BI

While observing articles in the literature review part of the study there came up use case examples from the banking industry. Moro et. al. offered the following explanation for the eager and early adaption for BI techniques and solutions: *"For bank firms to survive and even excel in today's turbulent business environment, bank managers need to have a continuous focus on solving challenging problems and exploiting opportunities"* (2015, p. 1314). They also found that decision-making supporting analytics, performance metrics, credit risk related evaluation, customer segmentation and retention are common fields for BI applications in banking industry.

IoT based BI applications have been applied in the banking industry also. They can enhance the bank's ability in scalability and lower latency in data collection. With IoT solutions gathered data can be even richer and provide better insights when used in analyses. (Ramalingam & Venkatesan, 2019, p. 2037) It seems that BI has been widely adapted in the banking industry and it has various utilisation purposes. Considering the amount of data that banks and financial institutions possess about their customers and surrounding economy these findings do not surprise.

3.3. Problems faced

In this subsection the common problems faced in the BI implementation are addressed more closely. It seems that in most use cases organisations deal with more than one challenge at a time. This can be seen to originate from the multidimensional nature of BI processes. (Stockinger, et al., 2015, p. 53) Next, the most common types of challenges that have been encountered in the BI process implementation are

summarized. Afterwards, a table is composed where the characteristics of the most common problems are presented.

One of the most common challenges that surfaced frequently from the reviewed articles was related to managing large amounts of data and various information sources. When systems are dealing with large amounts of data the lack of suitable techniques can cause problems. Suitable analytical methods must be used and there is a demand for better processing performance in these cases. (Ahmed, et al., 2015, p. 71; Stockinger, et al., 2015, p. 53; Gangadharan & Swami, 2004, p. 142; Bhatti, 2013, p. 26; Ramalingam & Venkatesan, 2019, pp. 2038-2039)

There were also multiple observations about the problems regarding the latency in acquired data that is collected from different sources in the organisation. Some of the more advanced web based IoT and BD techniques are also prone to connectivity issues. These factors hinder the development of BI processes towards real-time analytics. (Ahmed, et al., 2015, p. 71; Stockinger, et al., 2015, p. 53; Bhatti, 2013, p. 26)

One crucial step for a successful BI system is to discover and identify the information needs. Some cases faced problems in this phase and it could affect the analysis outcomes significantly. Data quality concerns are related to this observation also. It is important to ensure that the analysis outcomes are reliable and not misleading. BI system must provide an answer for the designated information needs. Otherwise it can weaken the confidence to make data-driven decisions in the organization. Enabling DDD requires data and BI supporting organisation culture. (Gyulai, et al., 2019, p. 37; Gangadharan & Swami, 2004, p. 142; Bhatti, 2013, p. 26; Brichini, et al., 2015, p. 9; Ally & Khan, 2016, p. 96)

Most common problems found	Typical characteristics of the problem
Managing large amounts of data and information sources.	- Data management and performance issues in data storage. - Problems in combining of various information formats from different sources. - Insufficient methods in data handling and transformation causes problems.
Mismatch in data frequency between information sources.	- Different latencies in acquired data from various systems. - Systems closer to real-time analytics are more prone to problems with delays in data delivery.
Identifying information needs and transforming produced information into actions.	- Problems in identifying the most relevant and crucial information needs. - Reliability of the results obtained with BI system. - DDD supporting organizational culture is needed to fully utilize the benefits.

Table 3 Characteristics of the most common problems

Above, in Table 3 the most common problems that came up in this literature review section of this study are summarized. Every one of these problems are addressed and typical characteristics of the problems described. These observations help in forming an answer for the second sub research question. This question was related to the typical BI processes being automated and what kind of challenges are met in the process. Also, it is good to keep in mind these challenges when making observations in the case study section. It can help one to make more detailed observations and pay attention to remarkable details in the process. Thereby it is time to conclude this literature review and move on to the next section which will explain the methodology used in the case study.

4. Methodology

By research methodology it is referred to the research strategy that has been chosen to approach and observe the studied topic. At the beginning of the research it is important to understand the differences between study methods and their applications. By choosing the best fitting research methodology it is possible to get the most out of the studied topic. Also, with the best fitting methodology for the study it is possible to combine the existing knowledge and research to the studied topic and try to get answers for the underlying research questions. Traditionally research methodologies are divided in quantitative and qualitative research methods. (Hirsijärvi, et al., 2009, pp. 132-135) In qualitative research methods the most common ways to collect research material are interviews, surveys, observations and analysing documented information. These methods can be combined, used side by side or treated on their own (Tuomi & Sarajärvi, 2018, p. 62; Hirsijärvi, et al., 2009, pp. 191-192).

Since automation of BI processes can be implemented in many ways there is no precise theory about it in academia. Qualitative research methods were chosen to gather more knowledge about the topic and possibly find more future research ideas about the studied subject. In qualitative research the study does not aim to make statistical conclusions. Moreover, the aim is to describe and understand the studied subject or phenomenon. (Hirsijärvi, et al., 2009, pp. 160-162) Therefore, qualitative research is usually used in unique and subject specific studies. The nature of conducted study is to examine the phenomenon in a real-world environment and form a comprehensive picture about it. The most important goal of the study is to bring up unexpected aspects of the studied topic rather than testing an existing theory. For further research and analysis purposes it is crucial to understand the context and setting of the study. (Hirsijärvi, et al., 2009, p. 164) Qualitative research methods are applied to research topics which aims to understand complex and abstract subjects. It is also applicable method to study topics that has not been studied yet or does not have a clear or unambiguous theoretical background. (Saarinen-Kauppinen & Puusniekka, 2006)

Qualitative research method fits this study well because the setting is unique and there does not exist one commonly theoretical framework for automating BI processes. Every company is a unique case with limitations and boundaries coming from in and outside of the company. These restraints can be linked to economical boundaries, resource allocation policies inside the organisation or derive from legislative norms that the company faces. It is reasonable to use qualitative research methods with exploratory nature in this study because this study aims to add knowledge about the topic and the main problem is hard to define. (Hirsijärvi, et al., 2009, pp. 129-131; Saarinen-Kauppinen & Puusniekka, 2006) Since the main problem is hard to define in this research topic identifying possible course of actions, gathering more knowledge and forming further hypothesis is reasonable way to approach the studied topic. (Yin, 2014, pp. 10-11)

In this kind of setting exploratory research has its advantages. In the beginning the studied problem can be quite broad but it narrows down while the research advance. If there surfaces new insights or evidence during the research the researcher has to be ready to react to this and even change the direction of the study. (Saunders, et al., 2016, p. 171; Ghauri & Grønhaug, 2005, p. 58) Even though (Côrte-Real, et al., 2014, p. 175) states in the article that BI and BA has emerged tremendous interest in academic and these are popular concepts this topic has not been studied holistically or unambiguously theoretically defined. In many academic papers the most commonly used terms BI, BA and BD are often used as synonyms and inconsistently. Taking these prevailing perceptions and conducted research into account it is justified to use qualitative research methods. In this study case study method is used to observe the real-world setting and try to see how it differs from the academical findings about the topic. One must keep in mind that by choosing this research method the outcome is case sensitive even though the topic is tried to be studied comprehensively (Hirsijärvi, et al., 2009, p. 161). It is just a narration about this case specific implementation of BI process automation and its applications.

4.1. Case study

Research strategies can be divided into three groups. These three groups can be described as follows: Experimental research, where it is measured how changes in

one variable effect on other variable(s). Survey research, where information is gathered in a standardised form from a group of subjects. Case study, where one individual or small group of phenomena is observed intensively and the surrounding setting is tried to be understood. (Hirsijärvi, et al., 2009, pp. 134-135) Case study research has a long history across different fields of academia because it is used for understanding and investigating studied phenomenon in relation to its historical, economic, technological, social and cultural context. Business studies have a long tradition of using real-world business cases as an example. It has its perks as being personal, accessible and down-to-earth way to present complex and unique problems. Case study can be seen as an exciting and educational representation of real-world business situation in a more appealing format for managers and decision-makers. (Eriksson & Kovalainen, 2008, pp. 115-116)

Insight gained from the case study are usually reported as a narrative of the experience and findings in studied economic and business context. Case study research also includes interpretations of the examined actions, events and processes. (Eriksson & Kovalainen, 2008, p. 115) Methodologically case studies are different from the quantitative, experimental and deductive research traditions in business research that aim to produce statistical conclusions and generalisations (Ghauri & Grønhaug, 2005, p. 171). Case studies are connected to the interpretative, ethnographic and field-research traditions. One must recognize that despite the qualitative nature of case study method quantitative data could be used to construct a case. Therefore, case study should be understood more as a research strategy or research approach rather than as a method. (Eriksson & Kovalainen, 2008, p. 116)

In this case study observations are used to gather information about the studied subject. With this method it is possible to get close to the real-world situation and get observations directly from the organisations and its processes. It has been argued that observations are suitable way to gather information from changing and complex environment. (Hirsijärvi, et al., 2009, p. 213) Studied implementation of BI process automation can change and evolve during the time. Therefore, it is justified to observe and analyse how it has been built at the moment.

Types of observation	
Systematic observing • Systematic and structured • Observations are made by external observer	**Participating observing** • Adapts according the situation • Observer participates in the situation

Table 4 Types of observation techniques after (Hirsijärvi, et al., 2009, p. 214)

Participating observing has been chosen to be used in this case study. This is usually used in field studies and where observer tries to understand the studied subject as a whole (Hirsijärvi, et al., 2009, p. 216). This helps to get more insights from the implementation of BI process automation in the studied corporation. When using observations to gather information for study one must bear in mind that researcher must keep these observations apart from his own observations and interpretations.

4.2. Research context

In this study the focus on one specific real-world example of BI implementation and examine how it has been built. In the theoretical background section of this study an outcome is faced where there does not exist unambiguous definitions and mutually agreed theory or model to build automated BI processes. This finding also supports the decision to choose case study as the research strategy because the implementation of a BI process automation in a real-world scenario is examined. It is interesting to see how it is implemented in production compared to the theoretical framework found from academia. Observations for the study come from inside the company. These internal sources can be such as observations how the process works, documents regarding the BI process and technical aspects of the implementation.

4.3. Reliability and validity

Scientific research aims to avoid misleading or incorrect interpretations. By using suitable research methodology it is possible to get more accurate observations and avoid misinterpretations. Even so, the reliability and validity of studies varies. When the reliability of the study is discussed it usually refers to the repeatability of the

research. Someone should be able to repeat the conducted research with exact procedures and get the same findings and conclusions if they follow the study exactly same way. Validity of the study comes from the ability of chosen research method to measure and observe the studied topic. The study can be seen valid if the chosen research method does not give misleading or incorrect interpretations and the study answers for the underlying research question. (Hirsijärvi, et al., 2009, pp. 231-232)

In qualitative research methods validity and reliability of the study can be hard to measure. Especially case studies are unique real-world settings that are prone to changes in surrounding environment. Many researchers have tried to reflect and explain in detail how they have gotten the result from study. (Hirsijärvi, et al., 2009, p. 232) In this study the reliability will be dealt with careful presentation of the case company and describing the observed BI process more in detail. Observations are described and made from documents regarding the BI process and examining the implementation itself. This study can be repeated and therefore the results can consider to be reliable. One notable thing considering the reliability of this study is that the observed BI implementation can change. These changes may come from business needs or technical changes in the process. Also, if the observations are not made objectively it can affect the conclusions and findings of the study. Therefore, observer must keep in mind that study related observations must be kept apart from their own observations and interpretations.

5. Case study

5.1. Presentation of case study target

In this study a Finnish bank which operates in the financial industry and employs directly over twelve thousand people is observed. The OP Financial Group as a company differs from normal companies in a way that it is a cooperative business. This means that the company is owned and democratically controlled by its members in equal proportions rather than normal companies that are owned by shareholders in portions of shares that they own or possess. There are already over two million owner-customers who own the OP cooperative banks and thereby the entire OP Financial Group. Customers are served in retail banking, corporate banking and insurance sectors by offering a wide range of banking, investment and insurance services. (OP Financial Group, 2020)

The financial industry has undergone fast changes under the digitalizing world. It has also profiled as an industry for innovation concerning IT and technologies to utilize data. These new innovations and technologies have enabled electronic communication channels, advanced analytical techniques to evaluate risk and BI applications to support decision-making. To stay on top of this development and survive in rapidly changing environment banks have to focus on solving challenging problems and exploiting opportunities. (Moro, et al., 2015, p. 1314) Financial industry serves an important role in the economy. It is desirable that banks could operate more efficiently and therefore contribute better on the overall economy.

Many banks are facing the growing regulatory burden and fintech start-ups trying to get their foot on different niche fields in financial industry. This is causing erosion to profit margins of many banks. There has been evidence of correlation between technological progress and productivity in banking. Banks have automated many of their basic operations and procedures which generates vast amounts of data. (Moro, et al., 2015, p. 1315) Having lots of data about customers and surrounding economic environment creates an excellent opportunity for banks to harness the benefits and possibilities of BI for business.

OP Group has three main business segments which are: Retail banking, Corporate banking and Insurance customers (OP Financial Group, 2020). This case has its focus on observing automation of BI processes in one specific department inside of OP Group. This department is named as OP Markets and it is part of the OP Corporate banking division. The mission of OP Markets is to sell and provide all background processes of Capital Markets Products to all customers across the OP group. Customers come from the OP cooperative banks, directly from OP Corporate bank or from OP Group's Baltic branch. Size of these clients vary from an individual retail customer all the way up to large institutional entities. Offered product portfolio consist of various instrument types such as: Equities, Interest rate- and FX derivatives, FX, Bonds, Structured products, Equity analysis, Equity and Debt Capital Markets services, Execution and advisory services and Custody services. (OP Markets, 2020a)

Currently at OP Markets BI is applied in sales reporting activities. Development of current system originates from demand to provide information and insights about sales figures, customers and the structure of revenue generation from operations. The process is built on top of a recognized need for this kind of information which helps to follow and lead sales activities more accurately. Management can follow the results in more detailed manner and take actions based on the insights obtained. (OP Markets, 2020b) In the following section the used research methodology is explained and then the study proceeds to observe the BI process more closely followed with inspecting the architecture behind this BI system.

5.2. Way of conducting the case study

Case study with observative approach has been chosen to be the research method used in this case study. More detailed presentation of this method and its characteristics is given in the previous section of this study. By participating observing it is possible to catch the finest details and insights from the implementation of BI process. This method was introduced in Section 4.1. and the technique is explained more carefully in Table 4 of the same section. The process itself will be observed in a real-world setting to see how it operates and runs. The observer in his role tries to adapt in the process and understand the way it works. The findings are compared in the light of the theoretical framework that has been presented and the found

characteristics of challenges compiled in the literature review part of this study. Also, internal documents of the BI process help to verify the findings and concentrate the focus to relevant parts of the process.

The structure of this case study builds up as follows. Firstly, it is observed how the implementation compares to the proposed general BI process model presented in Section 2.4. (see Figure 12). Secondly, it is observed how the current BI infrastructure compares to the typical BI architecture presented in Section 2.5. (see Figure 13). Both of these steps are ended with a summary. Afterwards, an overview of the BI implementation and recommendations for improvement are given. While participating in the process possible short comes are tried to be identified and notify differences compared to the theoretical framework. All these steps are performed to give recommendations for improvements and see if there are traces of similar challenges found and summarized in the literature review.

5.3. Current BI process implementation

When observing the current implementation typical features of BI process are clearly recognizable. The aim is to transform data and information into knowledge and intelligence with applying prior knowledge and intelligence to it as (Hannula & Pirttimäki, 2005) presented, see Figure 9 in Section 2.4. The nature of this process is continuous because it collects, stores, manipulates and produces information for the management in a repetitive and cyclic manner. Therefore, it seems that it is possible to apply and compare this process in the light of general BI process model which was proposed in the Section 2.4. For a short rewind see Figure 12. This proposed model consists of five steps in a following order: Defining information needs and sources, Collecting and processing the data, Analysing the data, Sharing the information and Actions and feedback.

The first step of BI process must be implemented with care in order to identify all of the relevant information sources. In this case the intelligence needs are derived directly from the purpose of this process. Therefore, trading systems, complementary manual sources and CRM system acts as the information sources in this process (OP Markets, 2020b). From these sources the essential information for such BI system can be

acquired. These data sources provide needed information for the currently used sales performance related analytics applications. Despite the current demands are met there can arise new information needs in the future as the process evolves. These needs can be classified into three categories as described by (Vuori, 2006), see Section 2.4. for more details about classifying information needs. Overall, it seems that the current implementation captures all of the needed data sources to answer the questions designated for the current BI implementation. Therefore, it can be noted that there does not exist a problem with identifying the most relevant and crucial information needs in the current implementation. Another problem related to this step that surfaced in the literature review section of this study was related to managing multiple information sources and information formats. When observing the current BI process there did not turn out data source and type management related issues.

Second step in this model incorporates the data collection and processing. Data from the previously described systems are collected and transformed with automated ETL pipelines. An exception to this makes the manual sources that holds complementary information for the process. (OP Markets, 2020b) In this step the data is gathered and unified into a format that it can be used in the BI process. Implemented automation helps the organisation to achieve results that are closer to real-time analytics. Current level of automation enables analysis and distribution of information from previous day records. One identified problem in the literature review section was a mismatch in data frequency between information sources. Latency in acquired data, especially in manual sources have caused some lag in reporting. It is recognized problem that systems closer to real-time analytics are more prone to disruptions with delivering data on time. Also, known problems related to handling large data masses can arise in the ETL processes too. It seems that there has been applied sufficient data handling methods in the organisation because there were not any major performance issues in this part of the process. Even though, there is always room for minor performance improvements and tweaking especially in the data transformation procedures.

The third step includes the data analysis phase where the collected and processed data is transferred into useful knowledge and information. BI tools are used in this step assisted with statistical and analytical methods. In the observed organisation a

commercial BI tool is used to create dashboard view over the transaction data. This provides a comprehensive overview of the sales activities and allows user to interact with it, slice and dice the data, and drill down into the transactions (OP Markets, 2020b). One-time ad-hoc analysis can be run on top of the processed data too. Also, there exist an interactive trend analysis tool that is built with Python programming language. These features are closely tied to the fourth step of the proposed general BI process model. This step contains the dissemination of information and insights to the users. Currently dissemination is built on top of a web-based portal containing links to different analytical tools and applications where the latest data is updated automatically. From the observations it seems that the BI process outcomes are communicated widely to the users. One remarkable observation was that multiple times the information users were interested more closely to the process itself and how the data has been acquired. This can be seen somewhat linked to the problem concerning reliability of the results obtained with BI systems. There seems to be a need for more detailed specifications of the underlying data and explaining the process to understand the results even better.

The fifth and last step is where the data-based actions are made by decision-makers. In addition to this, feedback from the process should be collected. Observations confirmed that the case company management is taking actions based on the results from the process as described in the documentation (OP Markets, 2020b). In this light it seems that the BI is useful and answers the questions it is designated to answer. Feedback collection from this process is not organized systemically in the observed organization. An example of systematic feedback collection would be questionnaires sent to the users. It was noticed that the team responsible for developing BI received feedback directly from the users and had dialogue according the needs and development ideas with the end users. Overall, it seems that there is no resistance against shifting towards DDD in the organisation.

To summarize these findings a table is composed where it can be compared how each of the different steps have been constructed in the observed setting. One column for actions taken or not taken as described in the general BI process model and other

column for recognized challenges that are related to the problems found in the literature review.

Step of the BI process	What actions have / have not been taken in similar manner.	Observed problems similar to those found in literature.
1. Defining information needs and sources	- Information needs and data sources have been defined. - There is no classification of the information needs into three categories as described in the literature.	- No data source management related issues were faced. - Open question regarding the information needs, could more information sources provide new insights?
2. Collecting and processing the data	- Automated ETL pipelines are used to collect, transform and store information from various sources as described in the model. - The process includes also manually loaded complementary information sources.	- Identified problems regarding the frequency of manual data sources hindering the BI process real-timeliness. - No major data handling performance related issues but room for minor tweaks.
3. Analysing the data	- Currently, two different BI tools have been used to transform data into knowledge and intelligence. - Indications found that insights and valuable information is generated.	- Not necessarily the reliability of results obtained with BI system but the way of presenting the outcomes is problematic occasionally.
4. Sharing the information	- Latest results are updated automatically and shared among the users via web portal.	- This links to the problem presented above. Information sharing may not be efficient if every user does not completely understand the results.
5. Actions and feedback	- Management uses the results to gain a comprehensive picture about sales and take actions based on the results. - Feedback and development ideas comes directly from the users.	- There was no challenges in transitioning from data to actions and the organisational culture seems to be in favour for data-driven decision-making.

Table 5 Summary of the findings about current BI process

5.4. Current BI architecture

Next, the infrastructure behind this system that enables the effective utilisation and automation of BI processes is being observed. As noted earlier in this study the development of BI processes closer to real-time analytics has gained more attention in the studies lately (Hellsten & Myllärniemi, 2019, p. 345). There is similar development in the observed BI system where information and insights are demanded rather sooner than later. This brings the focus to the underlying architecture how these systems are built. Differences in the current implementation from the BI architecture presented by (Rausch, et al., 2013) are being compared.

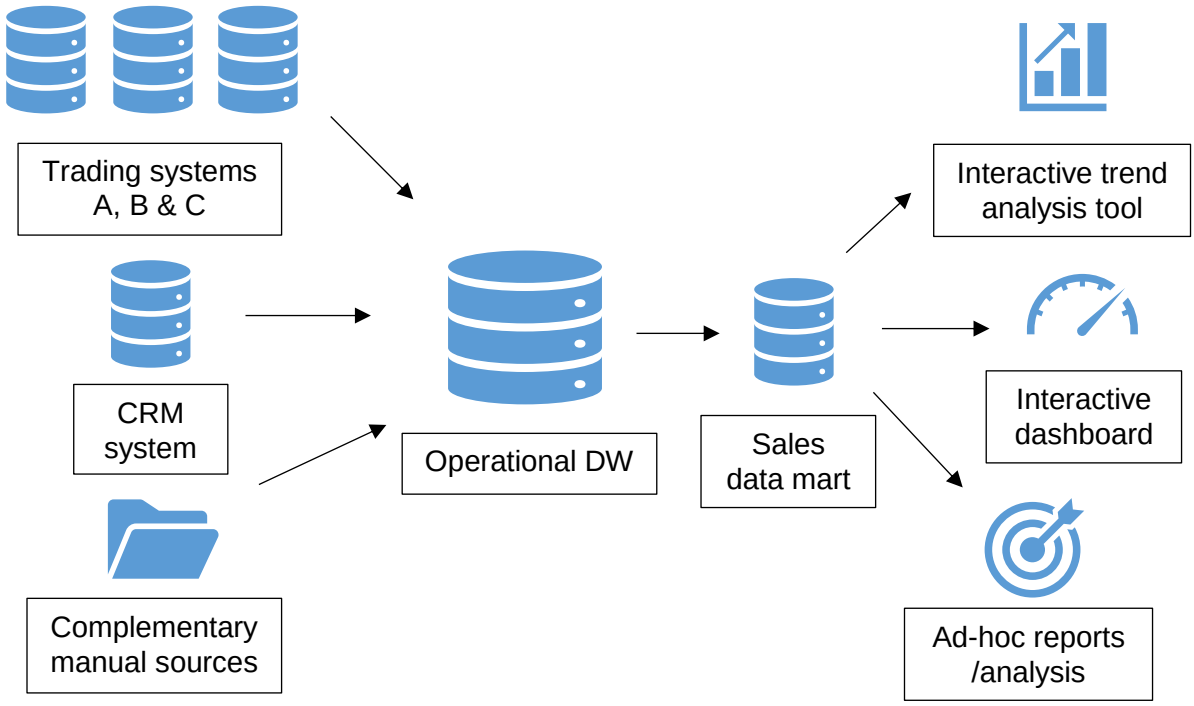

Figure 14 Current BI process architecture

This figure illustrates the current BI architecture of sales reporting activities in the organisation as described in the internal documentation (OP Markets, 2020b). There are clearly noticeable similarities to the BI architecture model discussed above. One can see that both examples collects structured and unstructured data to the data support layer where the information is stored. In the observed implementation the middle layer includes data manipulation, analysis and storing the outcomes as described in the literature. On top of these two layers an information access is built

with a described web portal. This provides unified platform to distribute the information to users with defined user roles and granted access rights.

As discussed about achieving real-time analytics in the chapter above one must find out the current state of data sources. In this case there are manual sources that hinder the transition closer to real-time analytics. Similar type problems with data frequency has been faced in the literature. Also, legacy systems and differences in the data formats must be considered in the BI architecture development. It seems that from the performance perspective the current implementation fulfil needs its needs. Based on the documented structure and the observations it can be concluded that the current architecture supports BI activities. Even though, there are recognized limitations in it that must be considered when planning further applications and development. To summarize the similarities and differences in the BI architecture layers the following table is composed. Own column is dedicated for the problems in current architecture that are related to the problems found in the literature review.

BI architecture layer	**What layers have / have not been implemented in similar manner.**	**Observed problems similar to those found in literature.**
Data support	- Data collected from various sources and formats have been done in a similar manner enabling its further utilisation.	- Legacy systems can cause troubles in data source management and development. - Manual data sources hinder real-timeliness of BI.
Information generation/ storage/ distribution	- Analysis layer includes procedures to transfer data into information in same manner as described in the literature.	- Current implementation uses suitable techniques and does not face any major data handling performance issues - Always room for minor performance tweaks.
Information access	- Results of the process are shared via unified web-based portal and this layer seems to be in line with the BI architecture model presented.	- Occasionally the results or underlying data have not been understood completely.

Table 6 Summary of the findings about current BI architecture

5.5. Analysis of current BI implementation

After observing the process and making the two comparative analysis it is possible to find out how the real-world implementation is different from the theoretical framework. Problems identified and faced in BI process automation were compiled and listed. These pinpoints can be used to address the faced problems in the organization and make the BI process more sophisticated and effective. It seems that some of the remarks in the literature share a common ground with the findings of this study. Also, the finding done by (Stockinger, et al., 2015) can be confirmed that organisations deal with multiple challenges simultaneously rather than just with one specific problem.

In addition to this, an evaluation of the sophistication of the observed BI system and ideas how to develop this process further are presented. In this case it seems that the process has been set up in a functioning manner and it produces insights that are valuable for the decision-makers. To evaluate the degree of intelligence produced in the observed organisation a scale presented by (Davenport & Harris, 2007) can be used. This scale was introduced earlier in the Section 2.3. (see Figure 8). Current BI system provides wide access and reporting on the sales data with ability to slice and dice the data and drill down to transactions. On the analytics features current system gives only statistical analysis about trends. At the moment the system does not provide forecasts of future trends, predictive modelling or analysis of different scenarios. In the light of these steps there is room for improvement if there is need for predicting different future scenarios.

Other development ideas are related to expanding the current system to answer more sophisticated questions. Such as identifying typical traits of the most profitable customers or predicting their behaviour. BI process could be enriched with more detailed data about the customers to understand their needs better. One possible way of conducting this type of analysis can be done by customer segmentation. Machine learning models using statistical methods such as self-organizing map (SOM) could be utilized in this process to help identify different customer types, clustering them and identifying the most profitable customers.

6. Conclusions

In this study the main objective was increasing knowledge about automated BI processes and its real-life applications by conducting a case study. This is done in order to see how the theory reflects into real-life applications and to bring up the challenges faced in this process. From the results it is possible to derive improvement ideas and direction for further development of the BI process.

The study begins with defining the theoretical framework by providing a working definition of the core concepts of BI. This was continued with a view of studies about the proven benefits of BI. Followed with examining the way that BI process works and a proposal for general BI process model is presented. Ultimately, the theoretical framework is concluded with a paragraph about BI infrastructure. After going through the relevant theoretical parts of the topic a literature review was conducted. This part of the study provides an outlook of BI use cases and problems faced in the process. Before moving to the case study the used research method is explained. By constructing the study in this way a profound ground for the case study is being built.

In the case study BI implementation in OP Markets, a department of a Finnish bank Corporate banking division has been observed. Comparison was made between the theory and practice of current BI process and its architecture. Observations about problems found in process were compared to the ones found in the literature review. Based on the remarks from earlier studies and the findings from the conducted case study improvement recommendations for the current BI process are given.

In this underlying last part of the study the findings are being discussed. Discoveries from different parts of the study help to format answers to the research questions. Also, the contribution and limitations of the research are being considered. At the end of this study possible future research suggestions are assessed.

6.1. Findings of the study

In this section the results of this study are being discussed and research questions answered. This section begins with addressing the sub research questions individually in the presented order. Followed with a summary of the problems faced in the case study and comparing these observations with the findings from literature review. After comparison the main research question is being answered. Now it is time to move to addressing the first sub research question.

What has been written about BI in the academic literature?

Matters regarding this subject were discussed in the second section of this study where the theoretical background was established. BI has been discussed widely in the literature starting from the 1950's and still there does not exist unambiguous definition for the term. Even though the amount of interest towards BI in the academic literature have experienced tremendous growth during the last decades. It seems that the term BI is considered as an umbrella term containing architecture, databases, tools, applications and methodologies. In the context of this thesis BI is understood as a systematic process of intelligence for obtaining data to serve the information needs of users and decision-makers to achieve insights and competitive advantage. The second sub research question derives its answer from the literature review part of this study and the question stands as follows.

What kind of BI processes are typically automated and what kind of problems are met in the process?

Remarks from the literature suggest that process that are repetitive in nature were the most common to be automated. Common characteristics of automated BI processes were the aspects of transforming data into actionable information and insights to support decision-making. All of these use cases had steps with collecting, processing, storing and analysing the data followed with disseminating the outcomes. These BI processes were usually applied to analyse transaction data, enhance production efficiency, reporting and monitoring of processes, risk and fraud analysis, provide a closer to real-time view of the organisation and producing more detailed information to

support the decision-making process. The most common problems faced in this process were related to managing large amounts of data and information sources. Challenges with mismatching frequency between different information sources caused trouble. And also identifying the information needs and transforming the produced information into actions lead to problems in the reviewed use cases. Third and last sub research question was closely tied to the case study section.

How can the findings compared to earlier research help on improving BI process automation?

In the case study BI process implementation was observed in a Finnish bank. Observations from a real-world scenario shed a light on the differences between the theory and real-life applications of BI process automation. It seems that the practice can benefit from the theory and vice versa. One can utilize the theoretical framework to evaluate the current BI implementation and make improvements based on the findings. This can bring up the possible problematic pinpoints that need more attention to make the process run smoother and become more efficient. Also, these findings can mark out the direction for developing the BI implementation in the future. Answers to these sub research questions helped to build up the path towards the main research question. Next, a comparison between the problems faced in the case study and ones found in the literature is composed.

Most common problems found in the literature review	**Observed problems in the case study**
Managing large amounts of data and information sources.	Legacy systems need to be taken into account in data source management and room for minor performance tweaks.
Mismatch in data frequency between information sources.	Manual data sources hinder the development closer to real-time analytics.
Identifying information needs and transforming produced information into actions.	Results are not communicated in detailed manner every time which disrupts turning produced information into actions.

Table 7 Comparison between the observed problems and ones found in the literature

This comparison underlines the fact that organisations deal with multiple problems at the same time rather than having just an individual problem. This was similar to the

discoveries made in the literature (Stockinger, et al., 2015, p. 53). Also, it is remarkable that how similar the observed problems are compared to the reported problems in the literature review. Of course, the variety of challenges that came up in the literature review was much wider. This is reasonable considering the extensive sample of use cases examined. By taking these remarks into account the main research question can be answered.

What kind of challenges are faced in BI process automation in the context of banking business?

Based on the findings in this study it seems that the problems faced in the banking industry are somewhat similar to the ones found in the literature. When considering the amount of data and information that banks possess it is not surprising that numerous legacy systems can cause trouble in managing data sources. Large amounts of data create a need for sufficient data handlings techniques which always possess room for minor improvements in performance. Real-timeliness of the BI process was disrupted due latency in manual data sources. This makes developing the system closer to real-time analytics considerably more challenging. Also, the effectiveness of turning produced information and insights into actions has room for improvement in some situations. Information must be communicated and presented in a way that everyone understands it.

6.2. Contribution of the research

This research contributes to the research area presented in Section 1.2. in numerous ways (see Figure 2). Theoretical background of BI was introduced and summarised to provide a comprehensive overview on the unambiguously defined term BI. At the end of the theoretical framework a proposal for general BI process model was given. It combined the different aspects and views of BI process from the literature. This model was used in the case study section of this study to evaluate the current implementation of BI process. Due the general nature of this BI process model it could be applied in other industries too.

As a result from the case study observations, ideas to improve the case company BI process were given. These ideas can be applicable to other organizations too. By

executing the same steps to evaluate the BI implementation in other organisations it is possible to gain their own ideas for improvement. And ultimately answers for the presented research questions were given.

6.3. Limitations of the conducted study

While considering the limitations of this study there are a few things to consider which are mainly caused by the context. This research conducted a case study in a specific department of a Finnish bank. Results derived from this one individual case are not necessarily generalizable. If the study tries to generalize the result, the sample size must consist of multiple observations. Also, one must keep in mind that BI can be used in various purposes and in this use case it was built for management and decision-makers to follow and manage sales activities. Every organisation is a unique setting and the problems faced may be different.

In addition to this there are some limitations considering the case study method. There always lies a possibility for not perceiving and capturing some of the finest details. This can leave out some essential information regarding the observed process. The underlying BI process is prone to changes and can evolve over time. Companies operating in the banking industry are recognized to be innovative due the high competition between peer financial institutions. The underlying BI process answers the needs of decision-makers that can change rapidly due the described factors. Also, while making the observations some of the problems may stay unnoticed and surface only under certain situations.

6.4. Suggestions for further research

Possible ideas for future research and fields of closer research came up along the process. One possible idea for future research would be conducting a larger scale survey. This could be done across the banking industry or financial sector companies to see what kind of challenges they have met in the BI process automation. It would result in wider knowledge about industry specific challenges in the use cases and possibly reveal common factors causing these problems.

Now the results observed come from one individual company. It would be interesting to see if these types of challenges are met widely in other organizations too. Similar findings would verify the remarks and conclusions made and make them more generalizable. Probably more detailed information would surface about the typical characteristics of the common problems.

The proposed general BI process model presented in Section 2.4 could be researched more in detail and tested in different setting. An interesting question arose while applying it to the case study, would it be applicable in other organisations too? The model seemed to correlate with the real-world implementation surprisingly well in the banking industry where companies tend to have lots of information about their customers. If the model is found to be working in other use cases too, it could provide benefits when planning or evaluating a BI process.

7. References

Ahmed, F. D., Jaber, A. N., Majid, M. B. A. & Ahmad, M. S., 2015. *Agent-based Big Data Analytics in retailing: A case study.* Kuantan, 2015 4th International Conference on Software Engineering and Computer Systems (ICSECS).

Ally, S. S. & Khan, N., 2016. *Data Warehouse and BI to Catalize Information Use in Health Sector for Decision Making: A Case Study.* Las Vegas, 2016 International Conference on Computational Science and Computational Intelligence (CSCI).

Arnott, D., Felix, L. & Yutong, S., 2017. Patterns of Business Intelligence Systems Use in Organizations. *Decision Support Systems,* Volume 97, pp. 58-68.

Azvine, B., Cui, Z. & Nauck, D., 2005. Towards real-time business intelligence. *BT Technology Journal,* 23(3), pp. 214-225.

Banda, M. & Ngassam, E. K., 2017. *A data management and analytic model for business intelligence applications.* Windhoek, 2017 IST-Africa Week Conference (IST-Africa).

Baryak, T., 2015. A Review of Business Analytics: A Business Enabler or Another Passing Fad. *Procedia - Social and Behavioral Sciences,* Volume 195, pp. 230-239.

Bhatti, N. D., 2013. Overcoming Data Challenges with Virtualization. *Business Intelligence Journal,* 18(4), pp. 25-31.

Božič, K. & Vlado, D., 2019. Business Intelligence and Analytics for Value Creation: The Role of Absorptive Capacity. *International Journal of Information Management,* Volume 46, pp. 93-103.

Brichini, M. et al., 2015. *Business intelligence for Business Intelligence: A case study at STMicroelectronics.* Athens, 2015 IEEE 9th International Conference on Research Challenges in Information Science (RCIS).

Brynjolfsson, E., Hitt, L. & Kim, H., 2011. Strength in Numbers: How does data-driven decision-making affect firm performance?. *International Conference on Information Systems,* Volume 1, pp. 541-558.

Chauldri, S., Dayal, U. & Narasayya, V., 2011. An overview of business intelligence technology. *Communications of the ACM,* 54(8), pp. 88-98.

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

Choo, C. W., 2002. *Information management for the intelligent organization: the art of scanning the environment.* 3rd ed. Medford: Information Today Inc.

Côrte-Real, N., Pedro, R. & Tiago, O., 2014. The Diffusion Stages of Business Intelligence & Analytics (BI&A): A Systematic Mapping Study. *Procedia Technology,* Volume 16, pp. 172-179.

Davenport, T., 2006. Competing on Analytics. *Harvard Business Review,* 84(1), pp. 98-107.

Davenport, T. H. & Dyché, J., 2013. Big Data in Big Companies. *International Institute for Analytics,* Volume 3, pp. 1-31.

Davenport, T. H. & Harris, J. G., 2007. *Competing on Analytics: The New Science of Winning.* Massachusetts: Harvard Business School Publishing.

Devlin, B. & Paul, M., 1988. An Architecture for a Business and Information System. *IBM Systems Journal,* 27(1), pp. 60-80.

Elbashir, M. Z., Collier, P. A. & Davern, M. J., 2008. Measuring the effects of business intelligence systems: The relationship between business process and organizational performance. *International Journal of Accounting Information Systems,* Volume 9, pp. 135-153.

Eriksson, P. & Kovalainen, A., 2008. *Introducing Qualitative Methods: Qualitative methods in business research.* London: SAGE Publications Ltd.

Gandomi, A. & Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management,* 35(2), pp. 137-144.

Gangadharan, G. R. & Swami, S. N., 2004. *Business intelligence systems: design and implementation strategies.* Cavtat, 26th International Conference on Information Technology Interfaces.

Ghauri, P. & Grønhaug, K., 2005. *Research methods in business studies.* 3rd ed. Harlow: Pearson.

Ghoshal, S. & Kim, S. K., 1986. Building Effective Intelligence Systems for Competitive Advantage. *Sloan Management Review,* 28(1), pp. 49-58.

Gilad, B. & Gilad, T., 1985. A Systems Approach to Business Intelligence. *Business Horizons,* 28(5), pp. 65-70.

Gilad, T. & Gilad, B., 1986. SMR Forum: Business Intelligence - The Quiet Revolution. *Sloan Management Review,* 27(4), pp. 53-61.

Gyulai, D., Bergmann, J., Gallina, V. & Gaal, A., 2019. *Towards a connected factory: Shop-floor data analytics in cyber-physical environments.* s.l., 7th CIRP Global Web Conference (CIRPe2019).

Hackney, D., 2000. Your Business Intelligence Arsenal. *Telephony,* 238(17), pp. 39-42.

Hannula, M. & Pirttimäki, V., 2005. A cube of business information. *Journal of Competitive Intelligence and Management,* 3(1), pp. 34-40.

Hays, C. L., 2004. What Wal-Mart Knows About Customers' Habits. *The New York Times*, 14 November.

Hellsten, P. & Myllärniemi, J., 2019. *Business Intelligence Process Model Revisited.* Vienna, 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2019.

Hirsijärvi, S. & Hurme, H., 2015. *Tutkimushaastattelu : teemahaastattelun teoria ja käytäntö.* Helsinki: Gaudeamus Helsinki University Press.

Hirsijärvi, S., Remes, P. & Sajavaara, P., 2009. *Tutki ja Kirjoita.* 15th ed. Helsinki: Tammi.

Hovi, A., Hervonen, H. & Koistinen, H., 2009. *Tietovarastot ja business intelligence.* Jyväskylä: WSOYpro.

Inmon, W. H., 2005. *Building the Data Warehouse.* 4th ed. Indianapolis: Wiley Publishing Inc.

Kim, K. & Bui, L., 2019. Learning from Hurricane Maria: Island ports and supply chain resilience. *International Journal of Disaster Risk Reduction,* Volume 39.

Laney, D., 2001. *3-D Data Management: Controlling Data Volume Velocity and Variety,* s.l.: META Group Research Note.

LaValle, S. et al., 2011. Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review,* 52(2), pp. 21-32.

Lönnqvist, A. & Pirttimäki, V., 2006. The Measurement of Business Intelligence. *Information Systems Management,* 23(1), pp. 32-40.

Luhn, H. P., 1958. A Business Intelligence System. *IBM Journal of Research and Development,* 2(4), pp. 314-319.

Marjanovic, O. & Dinter, B., 2017. *25+ Years of Business Intelligence and Analytics Minitrack at HICSS: A Text Mining Analysis.* Honolulu, University of Hawai'i at Manoa.

Moro, S., Paulo, C. & Paulo, R., 2015. Business Intelligence in Banking: A Literature Analysis From 2002 to 2013 Using Text Mining and Latent Dirichlet Allocation. *Expert Systems With Applications,* 42(3), pp. 1314-1324.

Moss, L. T. & Atre, S., 2003. *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications.* Boston: Addison-Wesley .

Nordic Institute of Business & Society NIBS, 2020. *Survey of large corporations 2020,* Helsinki: OP Financial Group.

Olszak, C., 2016. Toward Better Understanding and Use of Business Intelligence in Organizations. *Information Systems Management,* 33(2), pp. 105-123.

OP Financial Group, 2020. *OP Year 2019.* [Online]
Available at: https://www.op-year2019.fi/en/
[Accessed 5 March 2020].

OP Markets, 2020a. *Description of Markets business,* s.l.: Internal OP Corporate Bank report unpublished.

OP Markets, 2020b. *Sales reporting documentation,* s.l.: Internal OP Corporate Bank report unpublished.

Pellissier, R. & Nezhelele, T. E., 2013. Towards a universal competitive intelligence process model. *SA Journal of Information Management,* 15(2), pp. 1-7.

Ramalingam, H. & Venkatesan, V., 2019. *Conceptual analysis of Internet of Things use cases in Banking domain.* Kochi, TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON).

Rausch, P., Alaa F, S. & Aladdin, A., 2013. *Business Intelligence and Performance Management: Theory, Systems and Industrial Applications.* London: Springer.

Saarinen-Kauppinen, A. & Puusniekka, A., 2006. *KvaliMOTV - Menetelmäopetuksen tietovaranto [verkkojulkaisu].* [Online]
Available at: https://www.fsd.uta.fi/menetelmaopetus/
[Accessed 8 February 2020].

SAS Institute, 2018. *Lisää tuottavuutta S-ryhmän arvoketjuun.* [Online]
Available at: https://www.sas.com/fi_fi/customers/s-ryhma.html
[Accessed 25 February 2020].

Saunders, M., Lewis, P. & Thornhill, A., 2016. *Research methods for business students.* 7th ed. Harlow, Essex: Pearson Education.

Scholz, P. et al., 2010. *Benefits and Challenges of Business Intelligence Adoption in Small and Medium-Sized Enterprises.* s.l., European Conference on Information Systems (EICS).

Schroeck, M. et al., 2012. *Analytics: The real-world use of big data. How innovative enterprises extractvalue from uncertain data,* s.l.: IBM Institute for Business Value.

Spil, T. A. M., Stegwee, R. A. & Teitnik, C. J. A., 2002. *Business Intelligence in Healthcare Organizations.* Big Island, Proceedings of the 35th Hawaii International Conference on System Sciences.

Stockinger, K., Van Lingen, F. & Valente, M., 2015. Big Data Analytics in a Connected World. *Business Intelligence Journal,* 20(2), pp. 44-55.

Tan, K. H. et al., 2015. Harvesting Big Data to Enhance Supply Chain Innovation Capabilities: An Analytic Infrastructure Based On Deduction Graph. *International Journal of Production Economics,* 165(3), pp. 223-233.

Tedlow, R. S., 2003. *Giants of Enterprise: Seven Business Innovators and the Empires They Built.* New York, Harper Row.

Therieauf, R. J., 2001. *Effective business intelligence systems.* s.l.:Westport (CO): Quorum Books.

Tuomi, J. & Sarajärvi, A., 2018. *Laadullinen Tutkimus Ja Sisällönanalyysi.* Helsinki: Tammi.

Tyson, K. W. M., 1986. *Business Intelligence: Putting It All Together.* Lombard, IL: Leading Edge Publications.

Vuori, V., 2006. *Methods of Defining Business Information Needs.* s.l., Proceedings of the International Conference on Electronic Business (ICEB).

Williams, S. & Williams, N., 2004. Capturing ROI through Business-Centric BI Development Initiatives. *DM Review,* Volume 14, pp. 1-7.

Yin, R. K., 2014. *Case study research: Design and methods.* 5th ed. Los Angeles: Sage publications inc.