LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

School of Business and Management

Master's Degree Program in Strategic Finance and Business Analytics

Jaakko Salminen

SOLVING COMPLEXITY WITH VALUE-DRIVEN BUSINESS INTELLIGENCE

1st Examiner: Professor Eero Pätäri

2nd Examiner: Associate Professor Sheraz Ahmed

ABSTRACT

Author:	Jaakko Salminen
Title:	Solving Complexity with Value-Driven Business Intelligence
Faculty:	School of Business and Management
Major:	Master's Degree Program in Strategic Finance and Business Analytics
Year:	2017
Master's thesis:	Lappeenranta University of Technology, 71 pages, 18 figures
Examiners:	Professor Eero Pätäri
	Associate Professor Sheraz Ahmed
Keywords:	Business Intelligence, Business Analytics, Agile, Self-service

The objective of this study is to develop a business intelligence system for the case company that responds to the requirements of conducting high value creating business analysis. The research process starts with an observation period during which the current practices and systems in the case company are assessed. Previous literature suggests that focusing on timely delivery of insights, pervasive use and predictive analysis help create higher value. Methods and technologies suggested by literature to facilitate high value creation were put to a test using one of the most difficult topics in the case company.

Implementing the chosen methods and technologies provided significant improvement compared to previous experiences of the case company. Even so, the implementation of the system still required specialized expertise and thorough understanding of the data. Also the significant time used to study the situation in the case company helped to define the correct solutions to implement.

TIIVISTELMÄ

Tekijä:	Jaakko Salminen
Tutkielman nimi:	Solving Complexity with Value-Driven Business Intelligence
Tiedekunta:	Kauppatieteellinen tiedekunta
Pääaine:	Master's Degree Program in Strategic Finance and Business Analytics
Vuosi:	2017
Pro gradu-tutkielma:	Lappeenrannan teknillinen yliopisto, 71 sivua, 18 kuvaa
Tarkastajat:	Professori Eero Pätäri
	Tutkijaopettaja Sheraz Ahmed
Hakusanat:	Business Intelligence, Business Analytics, Agile, Self-service

Tämän tutkielman tavoitteena on kehittää case-yritykselle business intelligence alusta, joka vastaa korkeaa arvoa luovan business analyysin toteuttamisen vaatimuksiin. Tutkimusprosessi alkaa havaintojen keräämisellä jonka aikana arvioidaan case yrityksen nykyisiä käytäntöjä ja järjestelmiä. Aikaisemman kirjallisuuden mukaan tiedon nopea toimitus, jatkuva käyttö ja prediktiivinen analytiikka auttavat luomaan korkeampaa arvoa. Tutkimus testaa näiden tekijöiden implementointia yhteen case yrityksen vaikeimmista tapauksista.

Valittujen metodien ja teknologioiden hyödyntäminen johti merkittävään parannukseen verrattuna case yrityksen aikaisempiin kokemuksiin. Alustan kehittäminen vaatii kuitenkin erikoistunutta osaamista ja datan perusteellista ymmärrystä. Myös business analytiikan nykytilan arviointiin case-yrityksessä käytettiin merkittävästi aikaa joka helpotti oikeiden työtapojen ja teknologioiden valitsemisessa.

ACKNOWLEDGEMENTS

This thesis took a while to write, but it is finally ready. Thank you everyone for the patience. Special thank you to the case company for the opportunity, and support during the thesis project. I also want to express my gratitude to Sheraz Ahmed for guidance during my studies and the thesis project. Finally, I would like to thank my family for support during my studies, as well as my friends for their support, and the good times we had in Lappeenranta.

Helsinki, 16.11.2017

Jaakko Salminen

Table of Contents

1 Introduction

Due of digitalization human and business activities are generating increasing amounts of data that can be utilized to make better business decisions (Jank 2011). At the same time globalisation is driving competition in many industries and companies need to seize every opportunity to stay in the ever tightening race (Wiersema and Bowen 2008). As such, data does not facilitate better business decisions, but needs to go through a series of steps that transform it to knowledge that can be used in decision-making (Eckerson 2003). Business analytics is all about refining data into actionable insights, and investing in it has been commonly recognized to improve business performance (Tyagi 2003, LaValle et al. 2011, Kiron et al. 2014). However, the effective deployment of business analytics is not always straight forward. Analytics requires different technical and human capabilities and is largely dependent on the quality and quantity of the data that is available (Delen and Demirkan 2013a). While data can be employed to solve many problems it can also be notoriously difficult to extract value from (Acito and Khatri 2014).

Utilizing different information assets in decision-making is by no means new practice. In the past it might have been challenging to gather the needed data, but today, the main challenge is transforming the vast amount of data in to actionable insight (Olszak and Ziemba 2006). Despite the sufficiency of data volume it can still be difficult to access useful data, and sources are often unintegrated (Rasmussen et al. 2002). Business Intelligence (BI) systems are considered to be the solution that facilitates leveraging the data that has been encapsulated in different sources (Elbashir et al. 2011). BI systems consist of a set of integrated tools and technologies that are used to collect, integrate, analyse and make data available (Reinschmidt and Francoise 2000).

BI systems are the foundation for conducting business analytics and the primary goal of BI is to enable turning data into information (Pape 2016). The systems need to be able to facilitate timely delivery of insights and also take into account frequently changing business requirements (Larson and Chang 2016). With the highly dynamic needs of decision-makers (DeSarra 2012) and increasing need for business insights driven by predictive analytics it is obvious that current stovepipe implementations are unable to meet the demands of increasingly complex KPIs, metrics and dashboards (Delen and Demirkan 2013b). Agility

of the de-centralized information repository approach that most organizations have implemented is being constantly challenged (Zimmer et al. 2012). Traditional software focused development has been found problematic in BI projects. When it comes to BI, the value is not created by the software, but the information the system delivers. (Larson 2009) BI development requires close collaboration with business users, dynamic development methods and early delivery to ensure the creation of business value (Collier 2011; Larson and Chang 2016). The agile software development principles published in 2001 by Beck et al. call for more dynamic, less formal and customer focused development. This value-driven approach is a natural fit for BI development and practitioners have also applied it with success (Larson 2009).

The benefits of agile BI development have been recognized by practitioners and the success has been confirmed by a number of surveys. In addition, academic literature has picked up on the topic of agility in a BI context. Academics' definitions of agility are commonly based on Beck et al.'s (2001) manifesto and focus on the BI systems ability to rapidly respond when unexpected changes occur (Zimmer et al. 2012). The use of agile methods have been studied by the likes of Collier (2011) who argues that they improve the development process in both flexibility and time. His work is conclusive of both agile management methods and agile technical methods. Comprehensive literature reviews and analysis of agile principles in BI development have been carried out by Larson (2009), who lists best practices focusing on the information value chain, and Larson and Chang (2016), who include later trends of BI such as big data analytics as they propose an agile BI delivery framework. Also Krawatzeck et al. (2015) conducts a thorough analysis of related literature to develop a catalogue of the most suitable actions to undertake in agile BI projects. Baars and Hütter (2015) take a more practitioner-oriented approach as they develop their framework for identifying and classifying relevant measures of BI agility by interviewing experts in 25 companies. The starting point of their work is the division of BI agility introduced by Baars and Zimmer (2013), which classifies BI agility into content agility, functional agility and scale agility. On the other hand, Baars and Zimmer (2013) state that these agility types can be also categorized into the architectural layers of a BI system. Of the layers of the BI system the back-end data layer has been the most common topic of agile BI literature so far, highlights of which are such as Ambler's work on agile data warehouse development (i.e. Ambler 2003a; 2003b; 2012) as well as Hughes' publications (2008 and 2012). The popularity of the back-end as research topic is likely due to the fact that it accounts for more

than 50 per cent of the time and costs of BI projects (Davenport et al. 2001; Shen et al. 2012; Wang et al. 2012).

The focus that the data layer has received in the agile BI literature is by no means unjustified. However, agile development methods can be highly valuable in the development of the front-end applications as well. After all the front-end of the system delivers the data for analysis of the business user. Close collaboration and focus on business value in the development process should ensure the delivery of actionable intelligence to facilitate confident business decision-making. DeSarra (2012) has already applied agile methodology to the front-end development of a BI system in his case study, and Collier (2011) has also touched the subject, but the literature us still very thin. There are no case studies that would apply agile methodology and technology to development of BI system including front-end development which would include detailed descriptions of the development work. This type of research would have clear implications for practitioners, and would provide additional information to support previous findings of academics as well.

1.2 Research Objective

The objective of this study is to develop a business intelligence system for the case company that responds to the requirements of conducting high value creating business analysis. This primary research objective was formed in the research design phase, before the start of the actual research process. The research questions that would help reaching the research objective where formed in an iterative fashion during the research process. These questions are based on constant comparison of theory and the emerging empirical evidence to find the correct questions.

- What are the needs of the case company and what is limiting value creation of business analysis?
- How to build agility in the BI system?
- How to improve data integration?

1.3 Research Methodology

The value that this study will create, is not in testing of a theoretical model or a theory, but more related to issues in application of current knowledge in practice, in a specific context. For this purpose a case study methodology is applied. This methodology enables the research to be carried out in a specific setting using a variety of data sources (Baxter 2008). The research approach is more of the inductive type than the deductive type. The study will focus on inducting theory using a case study method. This includes within case analysis and, replication logic, that are distinctive to the inductive, case-oriented approach. (Eisenhardt 1989). In the deduction approach there is often a firm methodology that is followed to test a well-defined hypothesis, but in the inductive approach, as in this study, there is often no clear hypothesis that is tested. (Saunders et al. 2007) As McNiff and Whitehead (2000) point out, in this type of practitioner research, even the research question might not be clear before the start of the research process itself. In this case study the primary research objective was formed before the start of the research process, but the secondary supportive questions that would be required to get to the objective were formed during the research process.

Using a case study approach allows the use of both qualitative and quantitative evidence (Baxter 2008). In the study qualitative data is used to analyze the needs and issues of the case company related to conducting business analysis. The qualitative data is supported by quantitative evidence that focuses on the descriptive analysis of the data, and related meta data, that is used to conduct the business analysis in the case company. The primary data collection method for the qualitative data is unstructured participant observation, which is supported by documents. In participant observation the researcher may participate in the events that are being studied. The method allows the collection of information that might not be available in any other way, but is subject to the bias of researcher which should be carefully considered when conducting the research. (Tellis 1997) Participant observation is less used in management and business research, but common in sociology and anthropology. Yet, it can be a highly valuable tool in business research when used as the primary research method, and possibly supported by other methods. The main benefit of participant observation is a high level of immersion which allows the researcher to build an extremely deep level of understanding of the specific context. (Saunders et al. 2007) Another reason why observation was chosen as the primary data collection methods is that it enables a fresh

look at issues that otherwise would be taken for granted and thus not brought forth by other methods such as interviews or questionnaires (Robson 2002).

1.4. Structure of the Study

The thesis starts by presenting the important theoretical concepts related to the study. The theoretical section is divided to business analytics, and business intelligence systems and agile development. The first section will familiarize the reader with business analytics, what is it used for and what factors facilitate maximum value creation with business analytics. The purpose of this chapter is to provide an understanding of what is required from the business intelligence system which is used to conduct analysis. The business intelligence and agile development section presents the reader with the basics of business intelligence systems as well as agile principles. The most common business intelligence technologies are reviewed, as well as agile methods and technologies that can be used to develop agile business intelligence systems.

The empirical study begins after the theoretical section. First the case context is presented, followed by the findings made in the case company during the observation period. Next the development plan which was created based the observations and the theoretical framework is presented, and then the work done in each development iteration. The empirical part ends with an overview of the final business intelligence application presented from an end-user point of view.

The empirical part is followed by a discussion where the empirical findings are reviewed together with the findings of previous literature in order to compare them. The study ends with conclusions of what was studied and what kind of practical and academic implications the results have.

2. Business Analytics

Every time we make a phone call, use a customer loyalty card or search engine, we leave a digital footprint behind. These electronic transactions result in massive amounts of data that businesses are collecting. This data can be used to make strategic decisions, build better products and services and asses business risks. For most companies all of the new data has created significant opportunities and for some companies harnessing data has become a necessity. Data collection is especially important for companies that operate online, since they do not have the ability to meet their customers in person. Data collection and analysis is not only an opportunity for eBusiness, but since most business transactions are electronic also industrial companies can benefit from data analysis by optimizing their operations and cost efficiency. (Jank 2011) Many organizations are also trying to become more agile in response to globalized competition and rapidly changing market conditions (Delen and Demirkan 2012). To have the right information available at the right time for business decisions calls for renewed focus on analytics and new ways of working. Whereas, the amount of data and new analytical methods are an opportunity for businesses, they are a challenge for organizations that facilitate business analytics.

Business analytics is a topic of significant hype and the term is often misunderstood. There have been many attempts to define business analytics, but the task seems to be rather difficult. According to many definitions, the essence of business analytics is data-driven decision making, using statistical models and information technology to turn data into valuable insights (Cooper 2012, van Barneveld et al. 2012, Evans et al. 2012). Perhaps the most commonly used and a well-rounded definition comes from Davenport and Harris (2007) according to who business analytics means: “The extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions. The analytics may be input for human decisions or may drive fully automated decisions.” In general, there seems to be three aspects in the definitions that keep occurring: Data, use of statistical analysis, influence on decision making.

2.1 Different Types of Business Analytics

Business analytics facilitates reaching business objectives through analysis of trends and patterns in data, building predictive models to foresee opportunities and problems, and helping to find ideal solutions with optimization models. Delen and Demirkan (2012) provide a simple framework that divides these activities into three categories: descriptive, predictive and prescriptive, that are presented in figure 1. Descriptive analytics is more commonly known as business reporting. It includes standard, ad-hoc reporting as well as more dynamic data analysis, but it usually takes a backward looking view. Predictive analytics uses more advanced methods of statistical analysis and data mining techniques, and tries to gain a view of what will happen in the future. Whereas predictive analytics is trying to project the future conditions, prescriptive analytics is also trying to define the actions that are optimal for the situation. Algorithms and multi-criteria decision making models used in prescriptive analytics are beginning to bear resemblance to expert knowledge. Starting from the less advanced descriptive analytics, and moving towards prescriptive analytics we can see how the outcome shifts from less valuable standard reports, that merely support decision-making, to high value insight, that can already dictate the right actions. Effectiveness of analytics, no matter what type, depends largely on the quality and quantity of the data that is available (Delen and Demirkan 2012). To find meaningful insights from any data, majority of effort is spend on preparing, cleaning and standardizing the data (Kotu 2015).

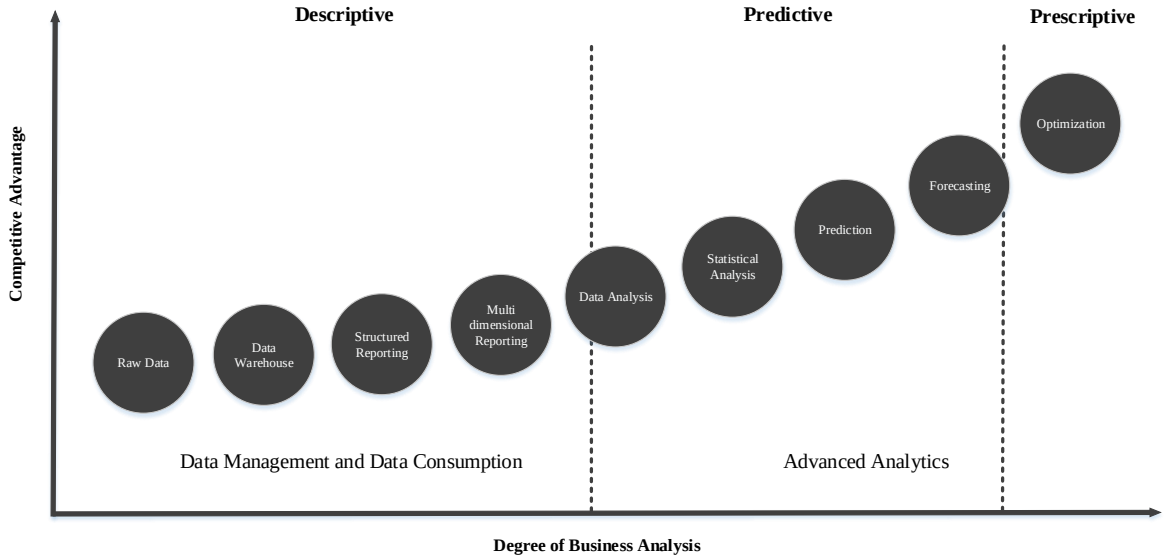

Figure 1. Three types of business analytics

Descriptive analytics provides information on trends and past and current events that give decision-makers context for future actions. Descriptive analytics is characterized by performance measurement and the use of key performance indicators. Since descriptive analytics also covers conventional business reporting it is the most widely used type of business analysis (Cook and Nagy 2014). In addition to standard, static business reporting descriptive analytics also includes the ability to drill-down to higher detail and to find out what are the actual causes for past and current performance, and where possible problems are. When combining data from different sources descriptive analytics can provide a comprehensive view on what has happened and why. This gives management context for decision-making. The role of descriptive analytics is to provide the historical information that decision-makers can react to what has happened. Descriptive analytics can also include automated alerts for potential problems. (IBM 2013)

Whereas descriptive analytics provides the past and current state of the business, predictive analytics is designed to help decision-makers plan ahead. The role of predictive analytics is to give decision-makers information about what could happen in the future and to anticipate likely scenarios. Using historical descriptive data, predictive analytics applies different types of models to predicted future events. Predictive analytics is characterized by the use of time series trends and correlations. By identifying patterns in the data, predictive analytics uses

advanced data mining and statistical analysis techniques to provide a fact-based foundation for decision-making. Simulation and forecasting models provide improved insights for decision makers and lay out the possible future states, so that a decision-maker can make informed decisions. (IBM 2013)

Predictive analytics can be seen as a subset of data-mining, which simply means finding useful patterns in data. In practice, data mining is searching large stores of data with sophisticated computational techniques. The aim is to discover structures such as patterns and trends which are meaningful and useful. (Kotu 2015) Since data-mining has been very computer science centered in the past, the basic data-mining concepts can be difficult to grasp for many business analysts. Fayyad et al. (1996) provide a clarifying overview of data-mining methods. The most common data-mining methods include techniques, such as simple linear regression, non-linear regression, logistic regression, decision threes and neural networks. These methods can be used in predictive analysis to classify data, find associations and clusters, and predict future values. The methods are not used solely for predictive tasks, but some of them can be used in descriptive analytics as well. (IBM 2013)

Prescriptive analytics is the most advanced type of business analytics. It explores possible actions and identifies optimal solutions from descriptive and predictive data. Using optimization and advanced mathematical techniques, predictive analytics suggest actions that have been identified based on a set rules, constraints and thresholds. The analysis should not only recommend a solution, but also reveal why it is recommended and what are the possible implications of the action, and also take uncertainty into account. Prescriptive analysis is especially useful when decisions are based on large and complex data sets that are impossible to understand without the help of technology. (IBM 2013)

According to a Eckerson (2003) 75 % of the tools that business analysts use are geared towards descriptive reporting tasks. These are usually standard reporting tools such as scorecards and dashboards that answer the question what has happened. 20 % of analytics tools are meant for deeper analysis, and answering why-questions. Examples of these analytics tools include spreadsheets and OLAP. Only 5 % of analytics tools are meant for predictive analysis, such as regression models, optimization and simulation. Eckerson (2003) does not include prescriptive analytics into the summary, so the share of prescriptive analytical tools is evidently very small, and these tools are for very specialized analytics.

2.2 The Purpose of Business Analytics

The idea of business analytics was born when tools and technologies were introduced that could process a larger quantity of information and recognize patterns in it more effectively than a humans can. When information technology progressed business analytics became more accessible and widely used. Zuboff (1985) noticed that when information technology is used to automate processes a lot of information about how an organization carries our it tasks is created simultaneously. Zuboff (1985) argued that this information could be used to build better understanding of the business and to also start improvement and innovation in production and delivery of products and services, which would improve the competitiveness of the business. In the first wave of business analytics the commercialization of enterprise information systems to a more generic format and key technologies such as the enterprise data warehouse brought more data available for decision-makers (Davenport 2013). The benefits of integrating organisational data from multiple data sources to a data warehouse have been discussed by author such as Anderson-Lehman et al. (2004), Goodhue et al. (2002), Watson et al. (2002), Wixom and Watson (2001) who also discuss factors that affect the successfulness of data warehousing projects. The new technologies and tools also required new competencies. Gradually business analytics started to require expert level understanding of information system, data management and analytical models as the systems became more advanced. Eventually business analytics became often seen as a service for business units, provided by analysts, while the infrastructure is provided by the IT department. Close co-ordination of these parties is still necessary for successful practice of business analytics. (Saxena and Srinivasan 2012)

The most recent development phase in business analytics was pioneered by the internet businesses and social network firms of Silicon Valley. These companies noticed that all the electronic transactions that took place generated vast amounts of data that had been previously overlooked. They began to analyse new kinds of information that quickly changed the role of data and analytics in those companies. Later the analysis of these extremely large amounts of information became known as big data. The data itself is characterized by the facts that it was not purely generated by the firm’s own transactions, but also acquired from external sources, and for a large part generated by human behaviour, meaning that a lot of the data was in an unstructured format. The enormous amounts of data

required new technologies for managing it. Since big data could not be efficiently stored and analysed in a single server, technology such as Hadoop and in-memory analytics were introduced. All of these changes required also a new set of competences from business analysts, that now needed to possess both computational and analytical skills. This lead to the creation of the role of a data scientist. (Davenport 2013) Even though the process of extracting valuable knowledge from large datasets, has been described already in the 1990s, by author such as Fayyad et al. (1996), Sasisekharan et al. (1996) and Simoudis et al. (1996) the practice has started to become more common among businesses only in the recent years.

It seems that business analytics was born not so much from a need than an opportunity created by advances in technology and changes in human behaviour. Chen et al. (2012) state that the role of business analytics is to help enterprises understand their business, market, and make better decisions. This rather general description applies for most cases, but in some industries the role of business analytics has become a central part of success. Examples of such applications are estimation of insurance prices, and optimization of logistics routes, as well as analysis of customer data to determine the effectiveness of a marketing campaigns (Kohavi et al 2002). In businesses that all customer contacts are purely online it is critical that companies have the ability to analyse the data generated by the customers when they visit the company website. Most companies do not have such specialised purpose for analytics as in the previous examples, but can still draw benefits from wide adaptation of analytics in decision making processes. According to LaValle et al. (2011) survey companies that substantially outperform industry peers have put analytics to use in the widest possible range of decisions. Thus utilizing all the opportunities that are available.

In order to use data as a basis of decision making it has to go through a number of refining steps, each requiring different technical and human capabilities. Eckerson (2003) sees that data is a raw material that needs to go through a refining process to be transformed into information products. As such, data does not facilitate better decision making, but needs to be transformed into information to begin with. Data is essentially discrete and objective facts regarding events. According to Joia (2000), in order for data to become information, it requires attributes, relevance and context, which all are provided by a human. Only a human mind can give data the purpose and relevance that transforms it to information. Of course the logic can be later build to an information system, and repeated automatically.

Figure 2. The transformation from data to action.

The next step in the refining process is to transfer the information into knowledge. Kock and McQueen (1998) define information as "descriptive and historical, relating primarily to the past and the present" and emphasize the predictive and associative nature of knowledge compared to information. They do however acknowledge that it is difficult to make the distinction between information and knowledge, since they are usually communicated in the same way. Davenport (1998) sees that knowledge is information combined with context, interpretation, experience and reflection. Emblemsvåg (2005) argues that business analytics cannot provide knowledge directly and that the outcomes of business analytics requires further work. Also the arguments promoting the predictive power of knowledge and involvement of humans in the creation of knowledge seem to be suggesting that humans have a major role in turning the results of business analytics in to knowledge that can lead to action. However, modern predictive models are starting to overtake the role of humans in providers of context and interpretation and eventually decision-makers. When the decision criteria are highly structured, data is high quality, and the decision making rules can be codified the decisions can be made with minimal human intervention by automated decision applications. Although, automated decision making processes can take care of repetitive operational tasks, more complex decisions are still left for humans. (Davenport 2010)

In order to use complicated analytical models the decision-maker needs expertise to understand what the models are suggesting. The numbers can tell many stories simultaneously and the reader needs to understand which ones matter. Not all need to have the capability to perform advanced data analysis, but the decision-makers should have a certain level of understanding. (Harris et al. 2010) According Davenport (2014) in order to make analytical decisions, one needs to understand the analytical process. As an example, the decision-maker needs to recognize that all analytical models are built on assumptions, understanding of which can be crucial to make correct decisions based on the model. Therefore, even managers need to have a certain level of analytical capability, even if they

do not create the models themselves. LaValle et al. (2011) see lack of understanding as one of the key obstacles in leveraging analytics for business value. It is difficult to make analytics based decisions if one does not understand what type of actions the analytics are proposing. Once the data has been transformed into information and then knowledge, it is vital to understand how the knowledge has been created so it can be finally turned in to valuable business decisions, and actions.

2.3 How to Maximize the Value of Business Analytics

Many surveys have found that companies that invest in business analytics are gaining competitive advantages and perform better than their peers. (Tyagi 2003, LaValle et al. 2011, Kiron et al. 2014,) Kiron et al. (2014) even describes analytic as “a common path to value”. It seems that there is no denying the value creating ability of effectively used business analytics. Sharma et al. (2014) argue however, that while there is evidence that investing in business analytics can create competitive advantage, the notion that business analytics leads to value needs some deeper analysis. They specify that the direct impact of business analytics is likely to be on decision making processes of the company, and that the improvement in the performance of the company are the outcome of the improved decision making processes.

Despite the extensive amount of analytics related literature published today and the long tradition on the area, there is surprisingly little attention paid to the non-technical factors in creating value with business analytics. Behavioural, organizational and strategic issues in maximizing the business impact of analytics are not receiving much attention. The focus has mainly been on how to get better data and better analytical models and tools for decision making. Focusing on improvements in discrete decision making situations that business analytics can enable, the focus on broader issues that facilitate an even bigger business impact has been neglected.

According to Sharma et al. (2014) how decision-making processes and resource allocations need to adjust to capture the full value of business analytics is of far greater importance than the current literature implies. There seems to be an assumption in the current literature that organizations can capture the value of business analytics while continuing to function as before. Same assumptions have characterized the introduction of Enterprise Systems and

Knowledge Management Systems as well. That organizations could capture the benefits, without major organization structure and process changes. Later research has recognised the criticality, or even necessity, of organizational and process adjustments to capture benefits from these systems. (Markus and Tanis 2000; Markus 2004; Kankanhalli et al. 2011)

Wixom et al. (2013) emphasize two factors that maximize the value of business analytics (figure 3). The first factor is speed to insight, which is concerned with how fast raw data is transformed into usable insight. Practices that facilitate greater speed to insight are automation, rapid identification of business requirements and reuse. First, automation, data integration and on boarding processes can cut the time of data preparation significantly. Second, agile development methods, such test environments and close co-operation with business users help to identify business requirements more rapidly, and deliver the right information faster. Third, approaches like data-as-a-service, or dashboard design catalogues will make it possible for companies to reuse resources and get information to business users more quickly. The second factor presented by Wixom et al. (2013) is pervasive use. By increasing the adoption rate of business analytics, companies can gain better return on their analytics investments. Wixom et al. (2013) present three drivers for companies to encourage more pervasive use of analytics. Frist, the user interfaces should be visually appealing. Companies should invest in proper dashboard design, because graphical visualization will make it easier for the end user to perceive the information. Second, delivering business analytics via mobile devices can also increase the adoption rate and increase productivity. Therefore, companies should make analysis accessible also through mobile devices. Third, companies can also promote user engagement through such practices as self-service analytics and gamification.

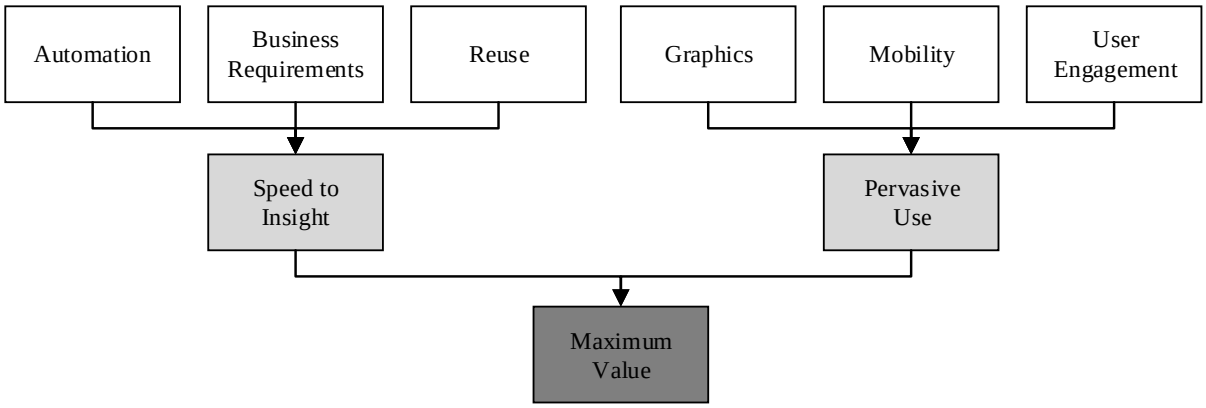

Figure 3. Factors that maximize the value of business analytics.

Also Hackathron (2004) promotes timeliness of insight as a driver of the business value created by analytics. Hackathron's (2004) value-time curve presented in figure 3 shows the relationship between time from a business event to the responding action and the business value that was created by it. This time from the event to action is known as action distance. It can be split into three components: data latency, analysis latency and decision latency. Data latency is the time required to capture, transform and store the needed data. Analysis latency is the time required to analyse and report the data to the correct person. Decision latency is the time required for the person to understand the situation, make the decision and implement it. Highest possible value creation can be achieved by minimizing each latency.

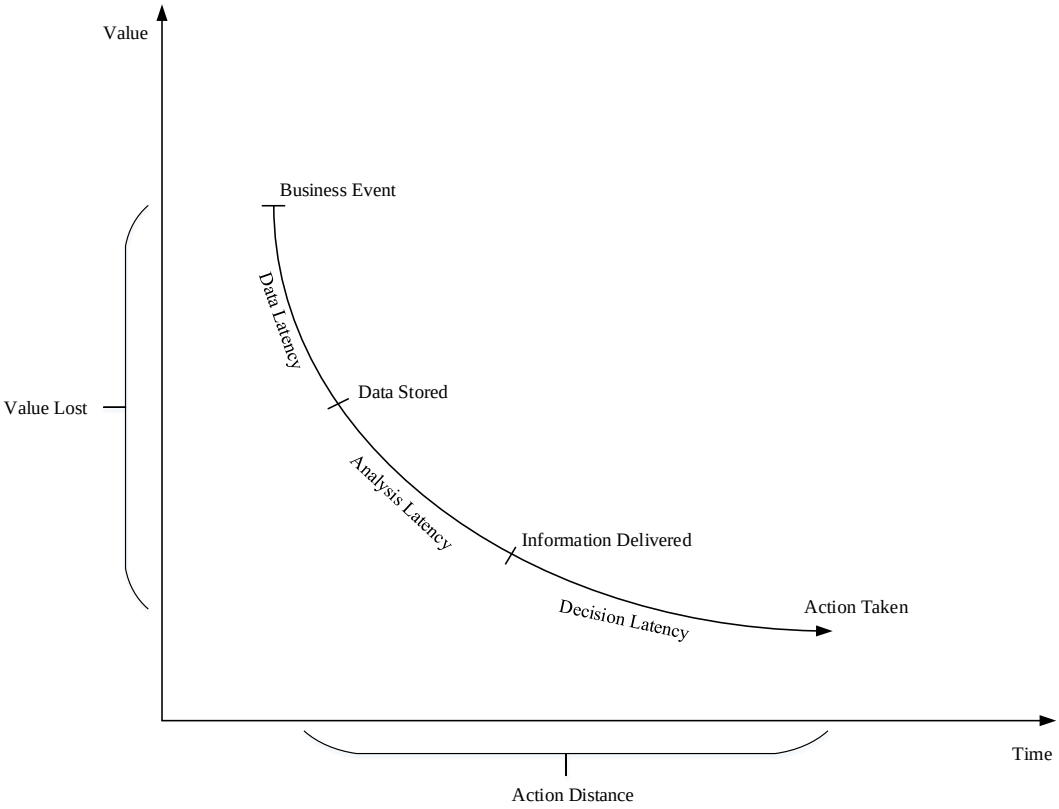

Figure 4. Action time curve

In practice, facilitating business analytics is carried out in collaboration of multiple parties that are responsible for different aspects such as IT infrastructure, analysis skills and business understanding. The use of business analytics is the most efficient when the special characteristics of different applications and methods can combined. According to Schläfkte et al. (2012) the combination of IT applications, management accounting applications and analytical methods will lead to the most efficient use of business analytics. It is important to understand the data when we are i.e. trying to identify causal relationships in the data. The identified factors and connections in the data will be enriched with more advanced analytical methods, and in reality, all this is possible because of the IT-applications.

3. Business Intelligence Systems and Agile Development

Decision-makers in corporations have always utilized different information assets in the decision-making process. In the past the needed information might have been difficult to come by, but now the abundance of data has made the task of turning the high volume of data in to knowledge the real challenge. (Olszak and Ziemba 2006) Many decision makers know the information they need is somewhere in the company, but they just can find it. Thus the decisions are made based on intuition and experience instead of data. According to Rasmussen et al. (2002) even corporations that have corporate information applications lack information needed for decision making for three simple reasons: The information in the applications is unintegrated, there is no historical data available, and data is not easy to get to even in the applications. According to Elbashir et al. (2011) Business Intelligence systems are considered as an innovation that will make it possible for companies to leverage the data that has been encapsulated in the enterprise systems

After organizations started to utilized computers and data processing also information system began to develop that supported management activities, such as planning, analysis and decision-making. Management Information Systems (MIS) of the 1960's were able to support management in routine management activities (Gorry and Morton 1971). Those systems were later developed to corporate Decision Support Systems (DSS) that were capable in supporting managers in unstructured decision-making tasks as well. Later in the 1980's and 1990's corporate information dashboards known as Executive Information Systems (EIS) became popular. The purpose of the dashboard was to give the key figures directly to the management in a simple user interface, without the need for assistants to write or print the reports. (Rasmussen et al. 2002) The idea was visionary, but before modern data warehousing and processing capabilities the solution was not practical. Current technology was not able to handle the large amounts of data spread across different source systems, which made the results incomplete and unreliable. In 1990's and early 2000's the advancement in data storage and processing capabilities made BI tools finally useful (Few 2013).

Business intelligence systems consist of a set of integrated tools and technologies that are used to collect, integrate, analyse and make data available (Reinschmidt and Francoise 2000).

Many studies, such as Ariav and Ginzberg (1985), Baars and Kemper (2008), Olszak and Ziemba (2006) see that business intelligence systems consist of three main components that are presented at figure 5. First a business intelligence system needs data, which is often spread in heterogeneous systems. The bottom layer of the system is the data layer, which consists of technologies that allow acquisition of data from various sources and storing it. The most common data layer technologies are enterprise data warehouses and ETL-tools. The second layer is the logic layer, which refers to technologies used for analysing the data. Technologies that offer effective analysis include i.e. OLAP and data mining. They can be used for descriptive analysis, such as discovering associations and exceptions, but also for predictive analysis, such as regression and time series analysis. The third component is the access layer, the business intelligence application which is used by the end user. This component combines the two previous components with a visual presentation that makes the data available for decision-makers to discover new insights. These three components all consist of multiple technologies, which means that the actual architecture is more complicated. In order to have a functional business intelligence architecture also tasks such as metadata maintenance and data quality control need to be included (Reinschmidt, and Francoise 2000).

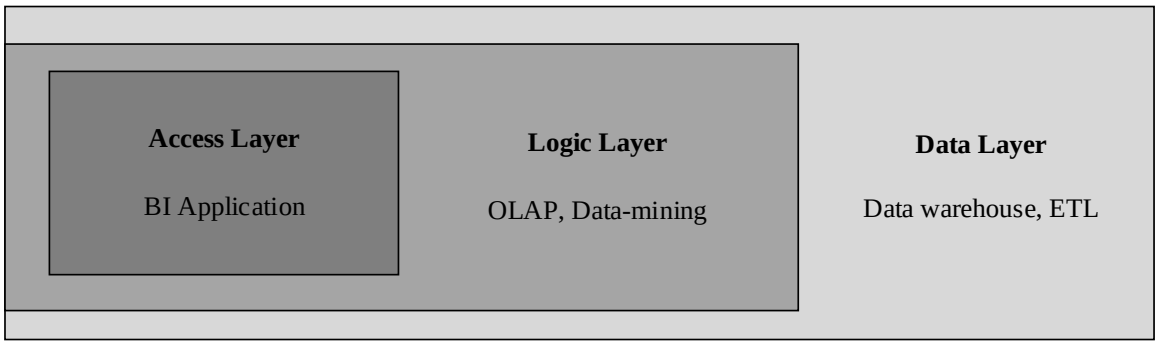

Figure 5. Business Intelligence system layers

Ong et al. (2011) have also introduced a multi-layer BI architecture, that includes a data source layer, ETL layer, data storage layer, end-user layer and metadata repository. Presented in figure 6. Unlike others Ong et al. (2011) include OLAP, and data-mining to the end-user layer, which illustrates the point that there is no common congruence between

researchers about in which layers different technologies exist. In reality there are often no clear boundaries between the layers, but the technologies are intertwined and implementations vary case by case.

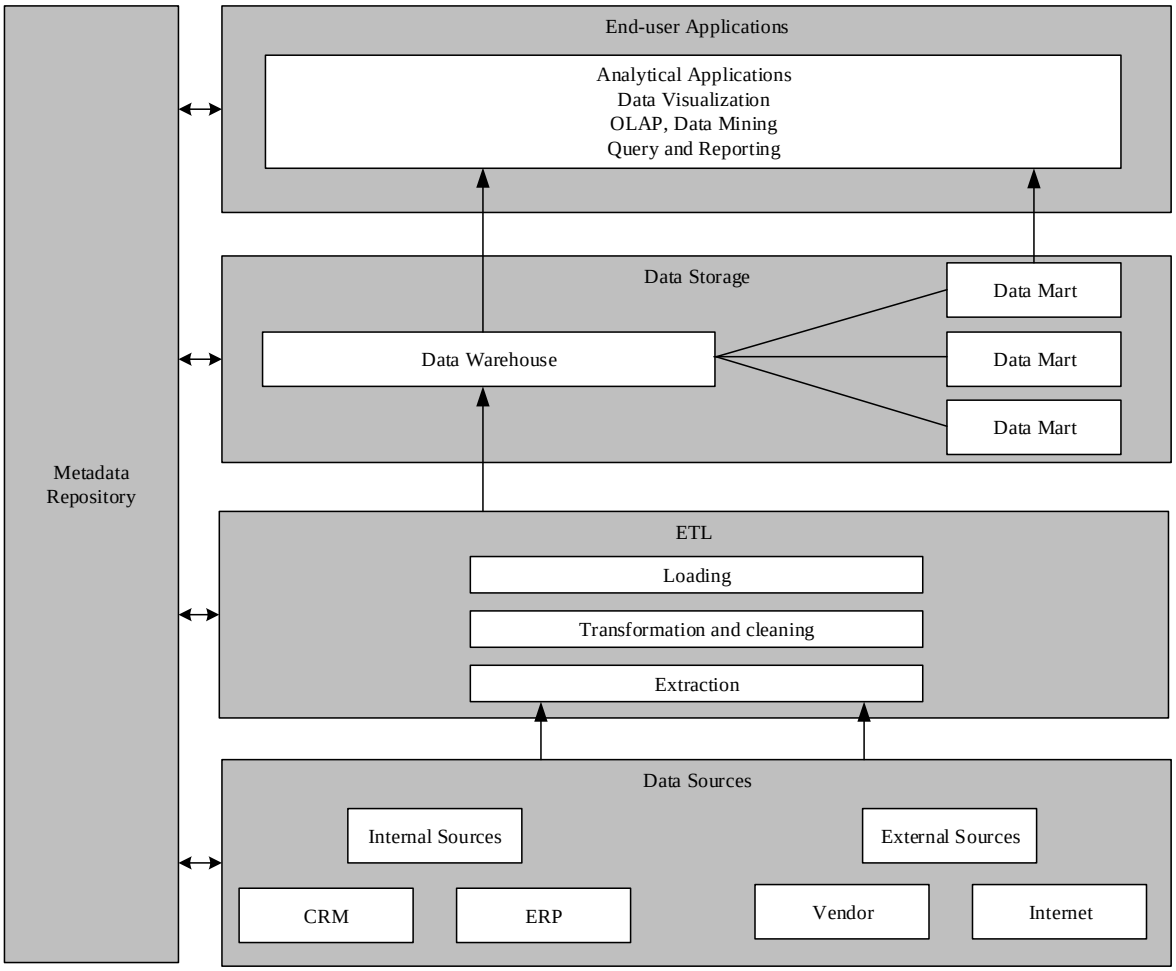

Figure 6. Multi-layered BI architecture

It depends on the purpose and the technology used that which components are emphasized in the business intelligence system. All of the technologies are quite easily misunderstood due to imprecise definitions and marketing hype. To help understand the role of the technologies Power (2001) has classified decision support systems to five categories based on the dominant technology component, including communications-driven, data-driven, document-driven, knowledge-driven and model-driven systems. Data-Driven decision

support systems emphasize availability of data. Access to large databases with structured internal and external data is the key. Model-driven decision support systems emphasize the potential of using statistical models, optimization, and other methods, which parameters can be modified by the decision-maker to help to analyse the situation. Knowledge-driven systems are focusing on finding hidden insights by utilizing intelligence data-mining methods. Document-driven system is a new type of system that uses unstructured data, gathered from web pages including text, images, video and sounds. These systems emphasize the possibility of searching information in the unstructured data. Communications-driven systems are meant to help groups in collaboration, and information sharing. The emphasis is on building an interactive system to facilitate problem solving and decision-making.

When planning a business intelligence project a company needs to carefully consider what is the purpose of the decision making support system. Design and implementation issues are different for communications-driven, data-driven, document-driven, knowledge-driven and model-driven systems. Since dominant technology component, target users and the primary deployment technology vary based on the purpose of the system it has to be considered what are the key issues that need to be solved. Model-driven systems primarily need to run the correct mathematical and analytical models, but the importance of the deployment technology is secondary. Conversely deployment technology is key for communications-driven systems. Decision support systems with multiple user groups can have complex deployment issues. Often the only possible deployment technology that allows collaboration and coordination is a web-based architecture, especially if the system is inter-organizational. (Power 2001)

3.1 Business Intelligence Technologies

Data has become an important asset for corporations, and at the same time managing the increasing volume has become more and more difficult. (Chaki 2005) Business intelligence systems are often mistaken for a pretty dashboard, but in fact BI systems consist of tools and technologies for extracting data from multiple sources; cleaning, transforming and integrating the data; loading it into a database; periodically refreshing the database; and

presenting the information to the front end of the system, which includes the dashboards used by the consumer. Chaudhuri and Dayal (1997) describe the tools for extracting, loading and storing as back end tools, and tools for querying and analysis as the front end client. The technologies can also be divided into three sections based on the BI infrastructure layers presented in the previous chapter.

3.1.1 Data Infrastructure Technologies

Transactional business systems are not meant for storing large quantities of data. From operational perspective, there is no benefit of storing a long history of transactions in the source system, but it can actually decrease the performance of the system. (Rainardi 2008) Operational data bases are meant for online transactional processing (OLTP), tasks that are a part of the day-to-day operations of a business. Operational systems and databases need to be able to execute repetitive, structured tasks which are usually isolated transactions. The key is to perform a maximal amount of transactions with consistency and reliability. Whereas the data in the transactional systems is used primarily for operational purposes, the data in a data warehouse is used for decision-making. The primary difference between a transactional data base and a data warehouse is the ability and capacity to store historical data (Rainadi 2008). The data warehouse is specifically targeted for analysing historical, summarized and consolidated data instead of isolated transactions.

Integration of data sources to a data warehouse requires three steps. First the required data is extracted from a source system. Then the data is cleaned and transformed to desired format, and finally loaded into the data warehouse. (Jörg and Dessloch 2009) This process is commonly known as ETL (Extract, Transform, Load). The goal of the ETL process is to transform data from legacy systems to a consistent format and load it to a data warehouse so that it can be used for analysis. Previously the ETL scripts had to be hand coded, which made development flexible, although the scripts were lengthy, difficult to document and change afterwards. To avoid the overhead caused by these issues, also graphical ETL tools were developed, that generate the code or function as engine-based tools. Especially the engine based ETL tools improved the ease of building ETL processes, because of the graphical interface. The shortcoming was the slowness of the row-by-row transformation of data

thorough the engine, which although has improved with later generations. (Zode 2007) However, as the ETL-process is the most time-consuming and complex part of a data warehouse project, all supporting methods are welcome (Petrovic et al. 2016).

In addition to designing the ETL-process also the structure of the data warehouse itself needs to be designed. Databases are generally structured as dimensional models, which store data into two types of data tables. Measures and metrics such as monetary values, quantities and KPIs are stored in fact tables, and descriptive fields such as department, product and unit of measure in dimension tables. This structure makes the data easier to understand and queries faster to perform. (Kimball 2013) Relational models are the most common design approach for not only dimensional databases, but all database design (Foster and Godbole 2016). Dimensional model implemented in relational database is known as star schema, for the star shaped structure (Kimball 2013). An example of start schema is presented in figure 7. The model withholds a single fact table, described by dimensions each of which are stored in their own table. For better performance and easier business queries, the dimensional model has also been implemented in multidimensional databases, which are referred to as OLAP cubes, for Online Analytical Processing technology used to process the data, and their cube shaped structure (Kimball 2013). An example of an OLAP cube is presented in figure 8.

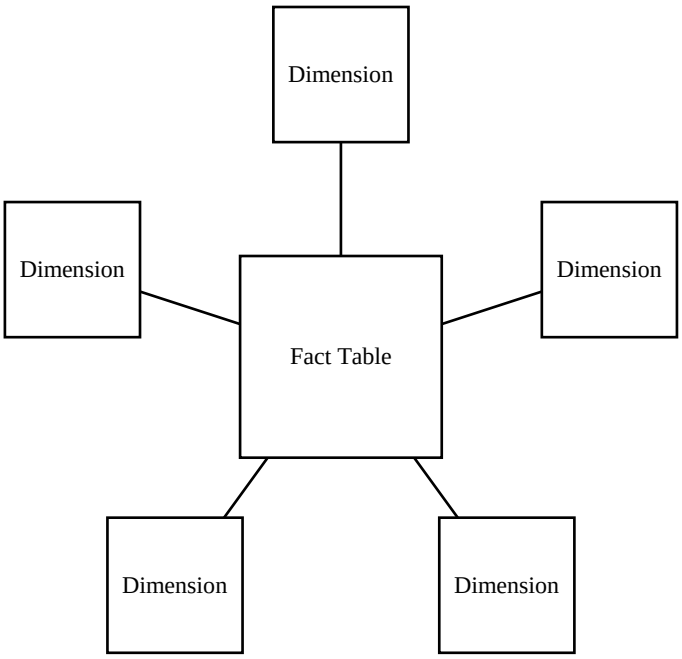

Figure 7. Snowflake data model

In addition to the relational model databases can also be structured in other more seldom used models. The hierarchical model is one of the earliest models and stores data in a tree-like structure, motivated by a common hierarchical organization structure (Jagadish et al. 1999). The network model is another early development, which has a similar three-like structure. Whereas, in the hierarchical model each node has a single parent, in the network model each node can have multiple parents. This makes a simple network model a more complex version of the hierarchical model. Either of these models are not used in new database implementations. (Foster and Godbole 2014). Object-oriented (OO) models are more recent database structures, that have been developed in order to avoid the mismatch between relational databases, and software applications, which are commonly programmed as objects (Alagić 2015). Whereas the relational model is based on entities by which the data is stored, the OO data in models are stored based on objects. An object in an OO model can be the same as an entity in a relational model, but an OO model object can also refer to in-memory objects. These objects can be for example a combination of attributes from multiple entities. (Ponniah 2003) Object based technology provides significant benefits when used

with software application programming, but may not be as advantageous as database design (Foster and Godbole 2014).

These data models are some of the most well-known, but there are also less used models and a number of combinations that can be used to implement a data warehouse. Besides the data warehouse there is a number of architectural alternatives enterprises can choose from. Often enterprises want to build an integrated enterprise data warehouse that gathers data from all areas of the enterprise, from customers, to assets and employees spanning also all geographical areas. These projects can be long and complex, require extensive models and take years to complete. Some enterprises are settling for a less complex solution known as data mart. Data marts are departmental subsets focused on a certain subject area, for example sales, or procurement. Data marts are faster to roll out, but might be difficult to integrate in case a need for an enterprise-wide model raises afterwards. (Chaudhuri and Dayal 1997) Another solution that has been increasing in popularity in the recent years is the data lake. As organizations struggle to keep up with the increasing volume, velocity and variety of data many have noticed that traditional data storage approaches are not able to cope. With traditional methods no data can be loaded before a data model an transformation framework have been build. This means that the data cannot be used before all data is in a structured format. However, with today's data volumes hardly all data is needed in structured format. Data lakes sift the traditional schema-on-write to a schema-on-read, which means that the transformation logic is created on an ad-hoc basis, for which ever purpose it is needed. When the data is stored, it is tagged so the relevant data can be made available for analysis when needed. Storing large amounts of data at low cost is made possible using technology known as a distributed file system, which also allows quick and easy scalability. (Jacobsohn and Delurey 2014)

3.1.2 Analytical Processing and Data-mining

The amount of data and number of databases in enterprise put pressure to develop analysis and interpretation techniques of the gathered data. Technologies and tools of intelligent and automatic data analysis eventually became a necessity to support business decision-making. (Rostek 2010) Early data solutions in enterprises are passive in nature, thus the data is

prepared for decision-making manually by analyst. More recent data solutions automate the analytical tasks that occur frequently which have well defined procedures. Implementing the analysis rules to an otherwise passive data warehouse mimics the work that an analyst does. (Thalmmer et al. 2001)

The logic layer enables analysis of data and distribution of the relevant information. The enabling analytical functionalities include Online analytical processing (OLAP), data mining, interactive reporting, performance management and ad-hoc analysis. (Baars and Kemper 2008) Multidimensional and often hieratical data models are commonly used to facilitate easy exploration of data. (Chaudhuri and Dayal 1997) OLAP first lined out by Codd et al. (1993) is the technology that has made multidimensional data analysis accessable for business users (Pedersen et al. 1999). Before relational databases and OLAP analysts had to often aggregate data using spreadsheets and pivot tables. Spreadsheets are useful for two-dimensional data, but when more dimensions are added, (i.e. time) they become unwieldy. Pre-aggreagating the data with database processing along multiple dimensions helps analysts gain insights trough a variety of views, without the need to do the aggregation themselves (Sarawagi et al. 1998). Multidimensional data models group data as facts that are numerical measures, and dimensions which are textual fields that charaztarize the facts (Pedersen and Jensen 2001). Figure 8 illustrase the multidimensional data model known as a data cube. The categorial attributes (dimensions) run along the sides of the cube, where as the numerical attributes (measueres) are in the center. In figure 8 the data model has three dimensions, product, country and date. These dimensions are associated with hierarchies that define the aggregation level, the granularity of the data. An analyst exploring the data cube has some simple operations at his disposal. Changing the level of aggregation in the hierarcy from lower granularity to higher (i.e. from country to city) is known as a drill-down, and aggreagting the detailed data back to higher granularity is known as a roll-up. The cube can also be rotated to show a particular face. This operation is known as pivoting. The cube can also be sliced to smaller subsets of the data. (Agrawal et al. 1997) The pre-aggregation makes it easier to analyse the data, but in addition storing the data in a this structural format makes standard business reporting more straightforward. Often business intelligence is only limited by the data that is available. If a data warehouse can provide structured, timely and clean data, generating valuable views of the business should be plain and easy. (Rasmussen et al. 2002) OLAP enables the creation of interactive reports accroding to the defined dimensions.

It optimizes the searching large data bases, and allows complex business analysis. (Olszak and Ziemba 2003)

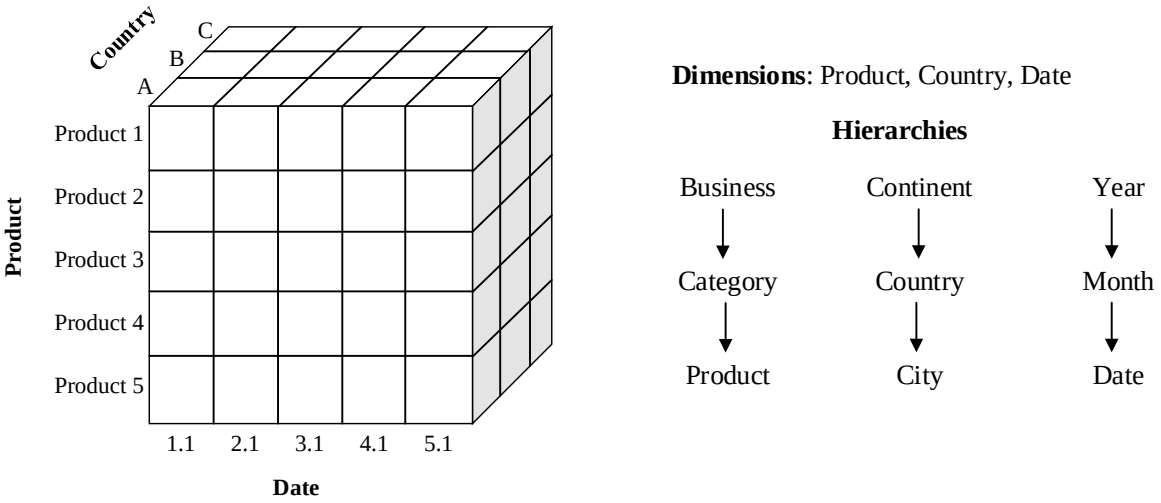

Figure 8. A multidimensional data model for sales data

The analytical benefits of OLAP based logical models are apparent. The data originating from the operational systems does not efficiently provide sufficient answers to business problems (Codd 1993). Pre-processing the data based on business requirements provides prepared data and analysis results in an understandable structure, thus significantly increasing the effectiveness of business decision-making. (Rostek 2010)

Sometimes solving business problems requires data analysis that is too difficult for humans to perform. This can be caused by volume of data or by lack of knowledge about what needs to be found in the data. Data mining methods can be used to discover patterns and relations in data. Automated data mining allows formulation of questions more abstrach than OLAP technology and discovery of new knowledge in the data. (Rostek 2010) Data mining is an important part of a process known as Knowledge Discovery in Databases (KDD), described by Fayyad et al. (1996) as the process of identifying valid, potentially useful, and ultimately understandable patterns in data. Data mining is a broad umbrealla term for a wide variety of methods that can be used in different domains to gain knowledge from data. Fayyaad et a.l (1996) classifies these methods to classification, regression, clustering, summarization, dependency modeling, and change and deviation detection. These methods have two primary

goals which are prediction and descripiton, and they can be used in two different scenarios. When the objecive is to find an answer to a known question, or to search for new, useful information without a predefined question. (Rostek 2010)

3.1.3 Business Intelligence Applications

The purpose of a BI application is to deliver information to the consumer in a user friendly format. Even if other layers of the BI infrastructure were well designed the system can still ineffective if the reporting layer does not deliver ease of use and present data in an intelligent way (Hovis 2002). (Kocakoc Erdem 2010) BI application provides access to the information for the end user who can use the application to monitor, explore and analyse it. A common way to monitor the information is a dashboard. (Few 2013) Corporate information dashboards, are among the major information technology trends that have developed during the past decade. When corporate dashboards were first developed they were known as Executive Information Systems (EIS). As the name implies the systems were built for top level executives to get a quick look at the most important KPI's and metrics. BI dashboards are becoming more common place as BI technology spreads across the corporate landscape and modern applications are designed to support decision-making in all management-levels (Olszak and Ziemba 2006). Like any high-tech solution also BI dashboards have been wrapped in a lot of hype by business intelligence vendors, and the role and meaning of dashboards is unclear for many companies (Few 2013).

Technology providers have varying definitions of a dashboard, but one of the most well rounded ones comes from Few (2013): "*A dashboard is a visual display of the most important information needed to achieve on or more objectives that has been consolidated on a single computer screen so it can be monitored at glance*" Visual display refers to emphasis of graphical presentation of data over text. The purpose of favouring graphics is to communicate the information more efficiently and comprehensively. In a way that the user can easily and accurately find the most important information. The content can be whatever is needed to achieve a specific task, not only high level information for executives. This content should all fit into a single screen, allowing it to be monitored at glance, without scrolling. This usually also means that the dashboard provides an overview, not the full

details. A dashboard must be able to quickly point out if something needs further attention, and possibly actions. More detailed information should be available easily and seamlessly by sifting to another view. Details can also be available by drilling-down in the same dashboard. However, this is not the primary purpose of a dashboard. (Few 2013)

By definition dashboards are meant for monitoring and do not include exploration capabilities. However, BI applications should allow the end-user to explore and analyse data beyond the predefined dashboards. (Few 2013) A BI application is the interface where a user is able to carry out the previously discussed multidimensional exploring operations. After the dashboard overviews help to identify the issues that need detailed analysis the user can drill-down to the details in another view of the same application and determine if actions or more information is needed. This process is known as performance monitoring. Few (2013) splits this process into four steps:

1. Update high-level situation awareness
2. Identify and focus on particular item that needs attention
    a. Update awareness of this item in greater detail
    b. Determine whether action is required
3. If action is required, access additional information that is needed, if any, to determine an appropriate response
4. Respond

The goal of this process is to monitor information in order to manage business performance. According to Few (2013) performance monitoring is based on the idea that to maintain or improve performance we need to be aware of the situation. Endsley (2000) sees that situation awareness has three levels. Fist, a person needs to have a perception of elements in the environment, then comprehension of the current situation, and last projection of future status. Few (2013) argues that BI dashboards should be designed to support situation awareness on all of these levels.

3.1 Agile Business Intelligence

Countless business decisions are still not based on analytics. This is surely not because of a lack of demand from decision makers. More often than not, the issue is the slow delivery of

insights. The conventional enterprise business intelligence architecture is still the backbone of business intelligence deployment, but there is a need to find alternative ways of delivering solutions that are able to provide faster time to value. (Imhoff and White 2011) Agile analytics, is a value-driven approach, which highest priority is early and continuous delivery (Collier 2011).

The big data phenomenon, increased volume, variety and velocity of data, has changed the business analytics and the ways information is used. This has also impacted the way business analytics and business intelligence applications are developed. The main consumers of business analytics are people whose jobs are not directly related to analytics, but use analytical tools to improve business performance. Business intelligence systems and other analytical applications are now becoming more accessible for business users, and being better integrated with transactional systems. This closes the loop between operations and analysis, resulting in faster analysis and quicker implementation of business actions. (Kohavi et al. 2002) The increased complexity in data and more dynamic business environment require faster, more dynamic, and more business oriented business intelligence development. Collier (2012) calls for practices that can be tailored to fit each environment depending on the dynamics of each project and organization. The principles of agile software development were first outlined by Beck et al. (2001) and have also been applied for business intelligence. The core ideas propose the development work to become less formal, customer focused, and more dynamic.

Collier (2012) states that the agile style of development is foremost, iterative, incremental and evolutionary. Instead of delivering the full solution at once, work is done in short iterations. The iterations are generally one to three weeks and never more than four weeks. The solution is built in small increments, each delivering a user-valued functionality. The solution evolves according to frequent user feedback. This ensures that the development work stays on the right track.

The goal of each iteration is to deliver value to the end user. While elegant models and complex data architecture or efficient ETL-scripts might be appreciated by the developer, the end user does not see any value in them. (Collier 2011) Access to, and presentation of information that helps solve business problems or make better decisions is in the interest of

business users. This is what every iteration must deliver business value. A model of the steps included in one iteration is presented in figure 9.

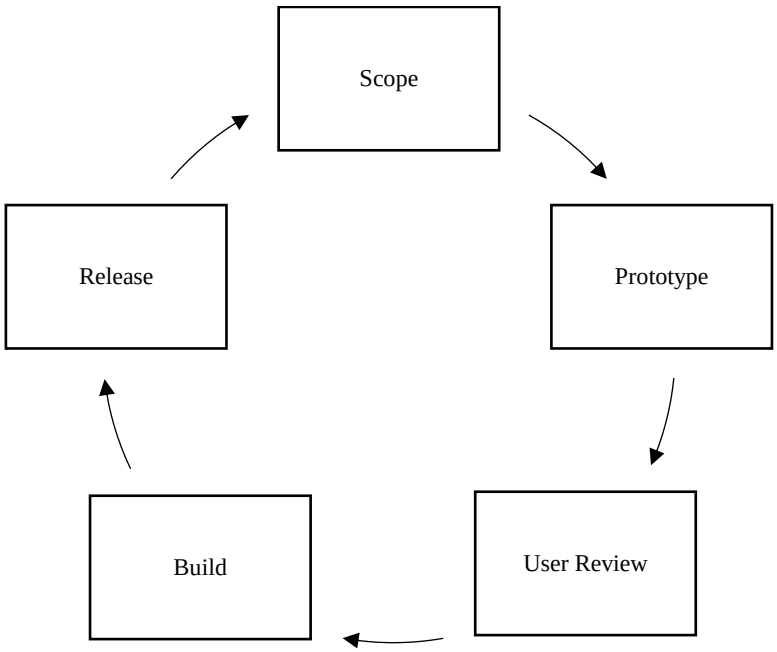

Figure 9. Development iteration steps (Desarra 2012)

In the traditional waterfall approach the testing and validation is made at the end of the development process, before delivery to the end user. Since agile is an iterative style of development, testing needs be a part of each iteration. Collier (2012) emphasizes including "rigorous testing early and continuously" in the development process. This ensures high production quality. The idea is to deliver working user features, not to produce hollow prototypes. A user feature is done when it is production ready and fully integrated into the working system.

"The only way to be truly Agile is to automate as many routine processes as possible" (Collier 2011). When routine tasks are automated the developers have more time to focus on delivering user valued features. Hence, truly agile teams seek to automate everything that is done more than once. Collier (2012) suggests testing as a critical routine process that should

not be done manually. Automated testing enables building new features frequently, always knowing how much time is needed for validating that everything is working properly.

In traditional project work the burden of ensuring that timelines and budgets are met falls too often solely on the development team. Agile methods seek for deeper collaboration between the development team and its stakeholders, and also acknowledges that the responsibility for success is shared. (Collier 2011) Agile project managers role is to hire the best people, give them the tools they need, then step aside and allow them to be successful. Project management is focused on facilitating a high degree of collaboration between the development team and a broader project community. The development is made in a self-organizing and self-managing team that decides itself what is made during each iteration, and holds itself accountable for reaching the goals. (Collier 2011)

The difference between agile analytics and a traditional development process is presented in figure 10. The upper line represents a traditional development process that is done in collaboration with the IT department. This approach is commonly referred to as the waterfall approach, because of the shape it takes in the graph. The process starts with the analyst and the business users laying down the requirements for the desired solution. The IT department does the development work, and testing. At this point nothing has been delivered to the end users (Y axis: remaining work to be delivered) which means there has been no business value created, and no feedback from the end users. Also all of the problems are raised by the end users after the solution has been finalized and is delivered to them. The solution is now build based on the requirements set in the beginning of the project. However, It is difficult to set the requirements, especially in today's big data environment. There are significant risks involved in the design phase when the requirements and the scope are defined. Unclear requirements, lack of understanding how the data is created and used, understanding the source systems constrains and unknown data quality are common obstacles in development projects. Deshpande and Desai (2015) add that learning and new discoveries can alter the project requirements, scope and priorities during the projects. At the point of delivery it might be difficult to make adjustments to the solution even when the business requirements are likely to change during the timely development process. According to Deshpande and Desai (2015) changing requirements and the difficulty to evolving them is constantly stressed as a major issue in data warehousing projects. Even when the requirements are well defined and the solution delivered perfectly according to the requirements, what was

considered ideal at the time of building the solution will change and the solution will need to be modified. This is clearly not the fault of the end users, but simply the nature of the dynamic business environment. Therefore, delivering the perfect solution cannot be considered a success unless required flexibility is also build in. (Bruni 2011) The aforementioned problems make the traditional waterfall development approach also prone to create lack of trust between the IT department and business stakeholders.

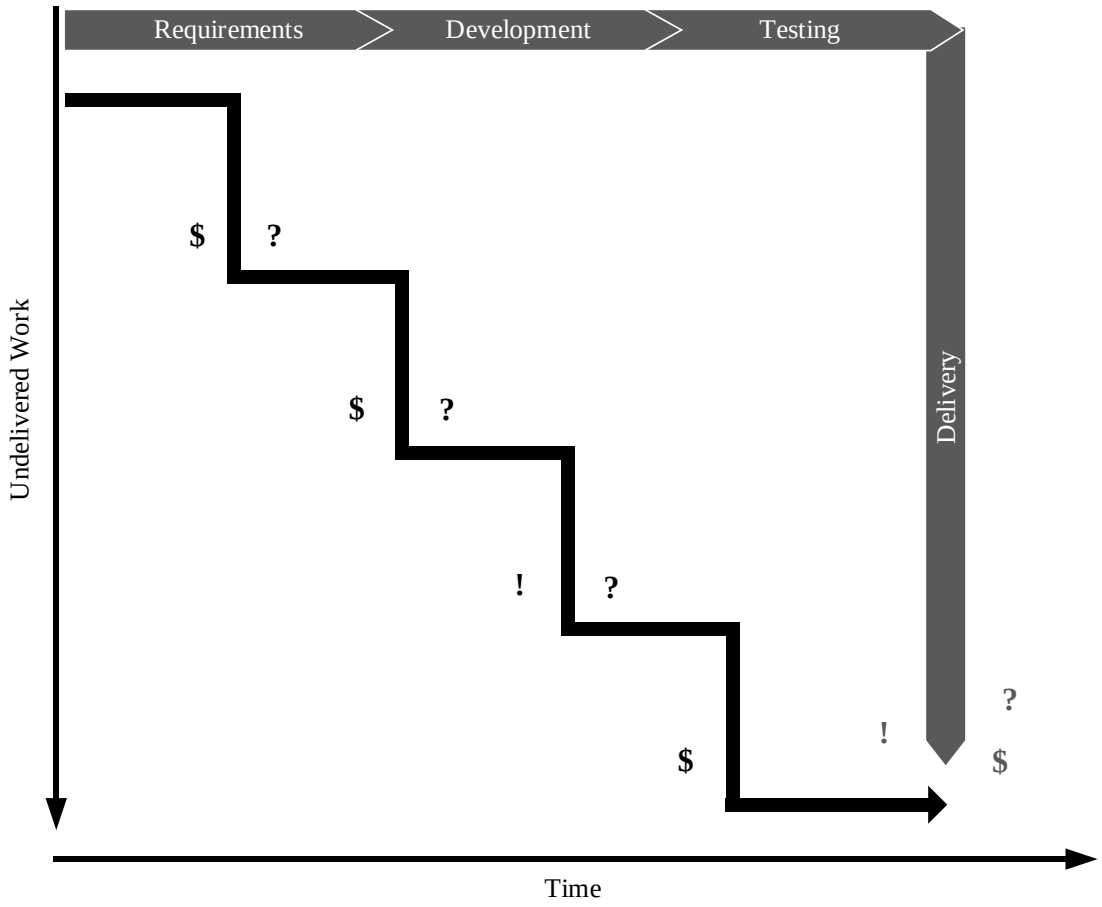

Figure 10. Traditional BI development vs. Agile BI (Lavastorm 2016)

The agile approach (the lower line in figure 10) seeks to deliver business value already in an early phase. Since the whole process is designed to be flexible and changes in the requirements are expected the process does not aim to deliver the whole solution as a big bang, but starts by asking what is the smallest, simplest thing we can do that creates business value. The development is then continued step by step, in small increments, each of which

bring business value and raise new questions or problems. The agile approach is focusing on delivering business value early, but also continuously thorough out the development lifecycle. The solution will evolve after each step, based on the frequent user feedback. Collier (2011) emphasizes agile analytics as an iterative, incremental and evolutionary process. Collier (2011) compares the agile development process to driving around in an unfamiliar city. You want to avoid driving too far without checking that you are on the right path. In agile analytics, the short iterations and close collaboration with the users ensure that the development stays on the right course. Yeoh and Koronios (2010) agree that the incremental approach makes managing risks easier and that it delivers measurable results that are visible to the client. Yeoh and Koronios (2010) also add that it improves the clients ability to take ownership of the solution, ease the transfer of knowledge and makes change management more efficient.

Muntean and Surcel (2013) propose a simple agile business intelligence framework that includes agile information structure, agile development, and agile business analytics. They argue that an agile business intelligence solution requires these three components in order to provide accurate information in the right format to the right person at the right time. Again the motivation is to build a business analytics model that is able to answer the rapidly changing business environment. Muntean and Surcel (2013) acknowledge the same issues with the traditional waterfall development approach and propose agile development methods that is very similar to other suggestions. In their framework, an agile information infrastructure is the basis of an agile business intelligence solution, the bottom layer, on top of which an agile analytical layer is built on. According to Muntean and Surcel (2013) the key in the agile information infrastructure is the ability to react to changing business requirements. It needs to extract and combine data from all data sources, external and internal, in any format. The proposed approach is data virtualization, to stop pre-transforming and pre-aggregating data and take care of these tasks on-demand. The data would be brought to the analytical layer thorough business views that integrate the data from different sources, while the source data remains in the original systems. The role of the agile analytical layer then would be to process the data using in-memory techniques, eliminating the need to pre-calculate in OLAP cubes or relational databases. The main benefits of the in-memory business intelligence solutions are quicker deployment, lower costs and ease of use for the end users. The solutions will also include advanced visualization features, such as drill-down capability and interactive dashboards. The agile analytical tools also support the agile

development process, allowing the business intelligence users to become less dependent on IT and become more agile trough the self-service solutions.

3.3 Self-service Business Intelligence

Business intelligence has a well-established business case and it has rooted itself into the daily operations of most companies (Gibson et al. 2004). However, Imhoff and White (2011) survey discovered that 78 % of business professionals need a faster time to value from business intelligence and analytics. This result builds the case for more agile approaches to facilitate business intelligence. One way to get faster time to value is to enable users to access data sets, queries and reports themselves, without any IT intervention. Relying heavily on the IT centric business intelligence support model is not sustainable anymore. IT departments cannot keep up with the frequently changing business requirements. According to Evelson (2011) Companies where IT addresses more than 20 % of business intelligence related requirements will see a snowball of a growing backlog of business intelligence requests. Enterprise IT systems, such as ERS and CRM systems can have a life cycle of several years, whereas a business intelligence application delivered in the morning, can be outdated in the afternoon.

The conventional waterfall development approach is not well suited for a majority of business intelligence applications. The real requirements often arise only after the end users sees something she can touch and play with. For most business intelligence applications the requirements that need to be addressed cannot simply be defined beforehand, since there is no way of knowing what type of analysis is needed tomorrow. This is why the user needs to be a major contributor in the development, and be enabled to adjust the application by herself.

One of the enablers of more efficient and more agile business intelligence environment is to enable users to become more independent, self-reliant and less dependent on the help of the IT department. An Imhoff and White (2011) study on self-service business intelligence lays out the following four focus points for facilitating an end user empowering business intelligence environment. (Imhoff and White 2011)

- Make it easy to access source data

- Make DW solutions fast to deploy and easy to manage
- Make BI tools easy to use
- Make BI results easy to consume and enhance

If you cannot access data nothing else really matters. You cannot make bricks without clay and you cannot do analytics without data. Data warehouse is the conventional IT-created business intelligence system component that enables users to access the needed data. In the self-service approach the all accessible data does not need to be in a data warehouse. Users should be provided the means to access external, operational and other data that is not in the data warehouse. (Imhoff and White 2011)

Self-service business intelligence calls for alternative deployment methods for data warehousing solutions such as agile methodology, SaaS solutions and cloud. There is no need to build an enterprise architecture, but business units can deploy their own self-service applications. User satisfaction is bound to increase when business users get a solution tailored to their own requirements, and the new technologies offer reduced costs as well as improved time to value. (Imhoff and White 2011)

One significant factor in the successful implementation of self-service business intelligence environment is the ease of using the applications. Fortunately business intelligence providers have focused on user friendliness for years. Reporting and simple analysis has been made quick and straight forward by the user friendly interfaces. The applications still need to meet the needs of more sophisticated analytics in order to fully achieve self-service. (Imhoff and White 2011)

Users have to have the ability to understand what the information presented to them means and how it is created. To promote adaptation self-service business intelligence must be easy to consume. The environment must be one where access, discovering and sharing reports, information and analytics has been made easy. This point is probably the most important one from the business users point of view. (Imhoff and White 2011)

3.4 Agile Data Integration

Decision support systems were built for supporting long term decisions in non-volatile environment, and are now facing demand to respect the increased market dynamics and frequently changing environment. Developing and maintaining new agile business intelligence systems poses huge challenges for IT organizations. The current business intelligence architectures in use are not fully applicable for modern agile ways of working, and agile development methods. They are rigid and inflexible and often unsuitable for readjustments, and in order for agile business intelligence systems to be created, the whole architecture must become more agile. (Knabke and Olbrich 2011)

As business requirements for analytics have evolved and density and velocity of data has increased the need for efficient ETL-processes and data integration has grown significantly (Theodorou et al. 2014). Many studies, such as Wixom et al. (2013), Muntean and Surcel (2013) promote agility as a driver to maximize the value of business analytics. However, according to Zaghloul et al. (2013) involvement of the IT department in business analytics not only increases costs, but also time delay. To achieve agile business analytics, business users need to have access to all data required to generate business analytics, with minimum involvement of the IT department.

The enterprise data warehouse is built for large scale, centralized, IT driven data modelling that can provide data for hundreds or thousands of users. A data warehouses can provide clean, structured data for frequent use, but require extensive amounts of time and money to build and maintain. The development of an ETL-process is the most expensive, time-consuming and complex part of a data warehouse project. (Petrovic et al. 2016). Companies are looking for more flexible and cost efficient ways of making their data available for business analytics and in recent years there have been alternative solution introduced for the enterprise data warehouse. Before an increase in computing power, the introduction of cloud computing and decrease in RAM memory costs, a centralized ETL-process and disk data storage was the best practice to prepare large amounts of business data for analysis. Today data integration and transformation can also take place in many different ways, in the cloud or in-memory.

When the user group of the business intelligence solution is limited, there is no need for a large scale enterprise data warehouse to facilitate the business intelligence application. In these cases the ETL-process can be made on on-demand fashion using tools that connect directly to source systems. Modern ETL-tools are able to acquire data directly from a wide variety source systems allowing business analysts far quicker access to data by eliminating the need to build an over-arching data solution. Once the needed data is loaded from the sources it is stored in RAM memory to avoid time consuming hard disk queries and allow faster processing (Garcia-Molina & Salem 1992). In-memory processing is presented in figure 11. The queries to the disk storage are carried out only once after which the data is processed in RAM memory. Winter et al. (2011) have discovered that application of in-memory technology can have a significant impact on how IT systems are used. It can turn static decision making into dynamic iterative analysis.

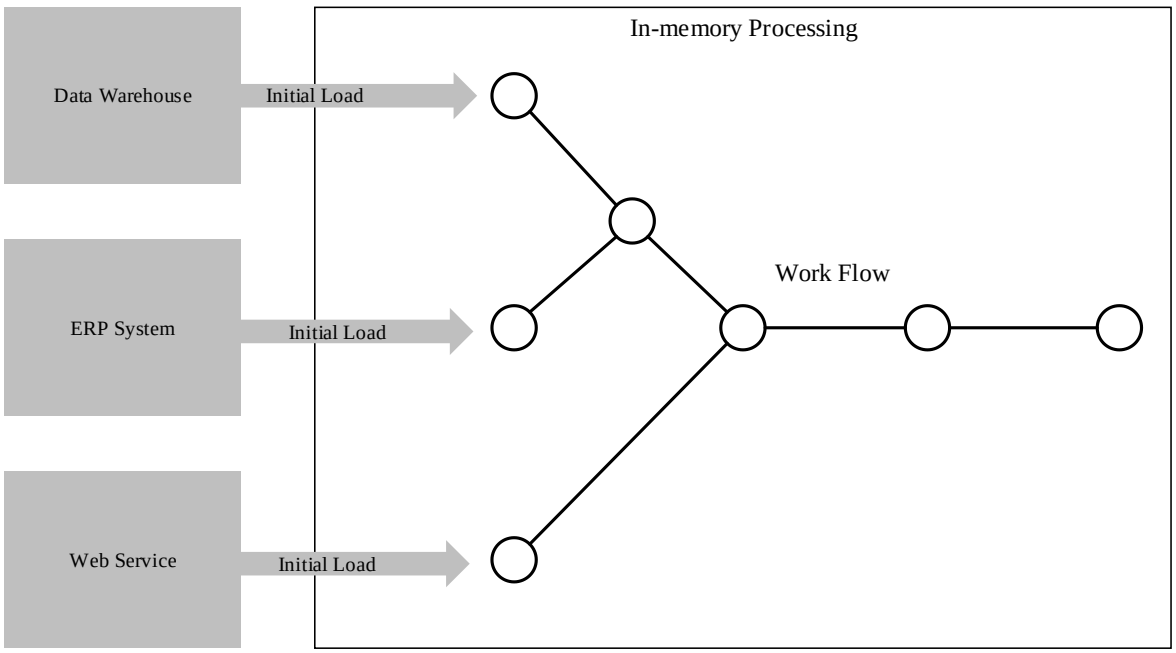

Figure 11. In-memory processing

Using an agile ETL-tool makes the integration of heterogeneous IT systems, and also other sources like spreadsheets and delimited files significantly more straight forward. There is no need to build a model to integrate whole data sets, but only the tables and columns that are

need. (Lavastorm 2014) When there is no integrated database available business analysts have traditionally had to rely on manual data preparation tools, such as spreadsheets, which require analysts to do a lot of manual work in order to build analytical datasets. This is a time consuming and unreliable way of working. Tools like spreadsheets are simply not robust enough to handle the data integration challenges and the amount of data and ad-hoc request of today's dynamic environment. (Alteryx 2015)

When data integration is carried out on-demand, the methods that are used need to be straightforward and quick to deploy. One of such method that has been gaining attention among practitioners is data blending. By definition data blending is simply the process of combining data from multiple sources to an actionable dataset for business decision making (Alteryx 2015). It allows the user to combine data from heterogeneous sources without any effort from IT to build a data integration architecture. According to Morton et al. (2012) this is exactly the novelty of data blending. The entire architecture supporting the data integration process is built on-demand and is adapt to evolving business requirements. Data blending is especially accustom to an agile approach as its key benefit is flexibility. The user has the freedom to control how the data is integrated and make changes on-the-fly (Morton et al. 2012).

In data blending the data tables from the secondary data sources are joined to the primary data source table based on the user defined keys. Using a single dataset makes the data model simple and easy to understand. The user can include only the data that is needed, which makes analysis of the data more straightforward. Since the process is simple, adding more data or modifying the model later on is not an issue. It is efficient to select only the required part of the data to a flexible dataset for analysis when data amounts are large, and in heterogeneous sources that could not be integrated to a data base without a substantial effort. A common data integration problem with heterogeneous data sources is granularity of data. With data blending these types of aggregation issues can be solved as they emerge, one table at a time. This makes integrating heterogeneous data much more efficient. After blending the created data set will be in the form of a flat data model. Compared to a common relational data base or the multidimensional model for OLAP presented above, the flat model is much more simple. In order to access or modify the data, the whole data set has to be read into memory. This makes creation of large data sets with data blending inefficient. Therefore,

data blending is not a scalable solution for building large data bases, but is effective in tackling specific, small scale integration challenges. (Morton et al. 2012)

4. Empirical Study

The empirical research process is carried out in three phases. These phases and the actions that were taken during them are presented in figure 12. The research process started with a data collection period, during which qualitative data regarding the case was gathered while actively reflecting on the theoretical framework to identify important findings. The qualitative data includes observation findings made during an observation period, as well as documentation of current reporting content, practices and IT system infrastructure. In the next phase this information was used to define the BI requirements of the case company. The requirements include both content related requirements and requirements related to the BI solution. In the last phase the BI solution was developed based on the requirements that were defined earlier. This phase began by planning the development work and discovering possible data sources and limitations in the data. Theoretical literature was also heavily utilized in the planning phase. After the initial plans had been drawn up the actual development work was carried that tested the agile methodologies in action.

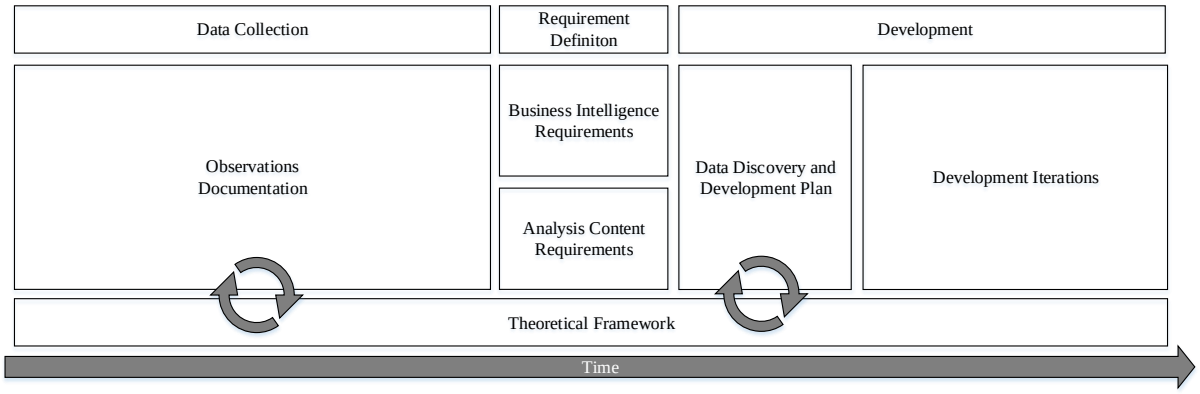

Figure 12. Research process

Because the observations were made during an extended period of time the reliability of the study is not threatened by participant error nor bias. If the data was gathered with a single interview or questionnaire the threat of getting unreliable data would be higher. The possible variation in the data provided by the subjects is balanced by the extended observation period

and the diversity of different situations when data is collected. The threat posed to the reliability of the study by the researcher is higher. The main threat to the reliability of this observation study is observer bias. The fact that it is not possible for the observer to fully detach himself from his own perceptions that impact interpretation cannot be avoided. In order to control observer bias all conclusions made during the data collection were thoroughly scrutinized.

There are no significant threats to the internal validity of the study. The ecological validity of the study is high, meaning that the findings should be well valid in this particular context. Especially considering the extensive time period of the study. However, the external validity of the findings should be considered more thoroughly. To extent to which the result are generalizable to other research settings is difficult to estimate. The findings can be assumed to be applicable to settings that share the main characteristics of this case.

4.1 Case Context

The case company is a large global manufacturing company. The company is involved in multiple different business areas some of which are further divided into business units. The company structure is fairly complex, which complicates enterprise wide reporting. It is also common in the company for different business units to use separate IT systems and not build enterprise wide systems. This approach has its benefits on operational level, but surely makes group wide analysis more problematic. Until recently the company has focused on descriptive reporting, but is now pushing towards more analytical decision-making and implementation of predictive analysis. This new way of working means that the reporting requirements are bound to change frequently as they are still finding their place. These three characteristics of the company: complex business structure, heterogeneous IT infrastructure, and frequently changing requirements form the unit of analysis of the case study, the results of which are applicable to similar contexts. Next these characteristics of the case company will be viewed in more detail.

The first characteristic is the case company’s hierarchical business area structure, and there are few additional factors that make this case environment problematic from a reporting perspective. The businesses conduct intercompany trade, which complicates consolidation

of data to a higher level in the entity hierarchy. Number of production units operated by a business unit ranges from few units to tens of units and operations in some production units are split between multiple business unit. These issues will need to be considered carefully when building systems for enterprise wide reporting and analysis. However, the BI infrastructure should not be made too complex and difficult to adjust, since business structures are bound to change from time to time.

The second interesting characteristic of the case environment is the heterogeneousity of the IT-infrastructure. The needed data for variable cost analysis is not available in a centralized system or a database. The data lies silos of each business's systems, and a number different systems needs to be integrated to get an enterprise-wide view. The case company's IT infrastructure is outlined in figure 13. Naturally, enterprise-wide financial accounting systems are in place, but integrating the other required data sources will be a challenge.

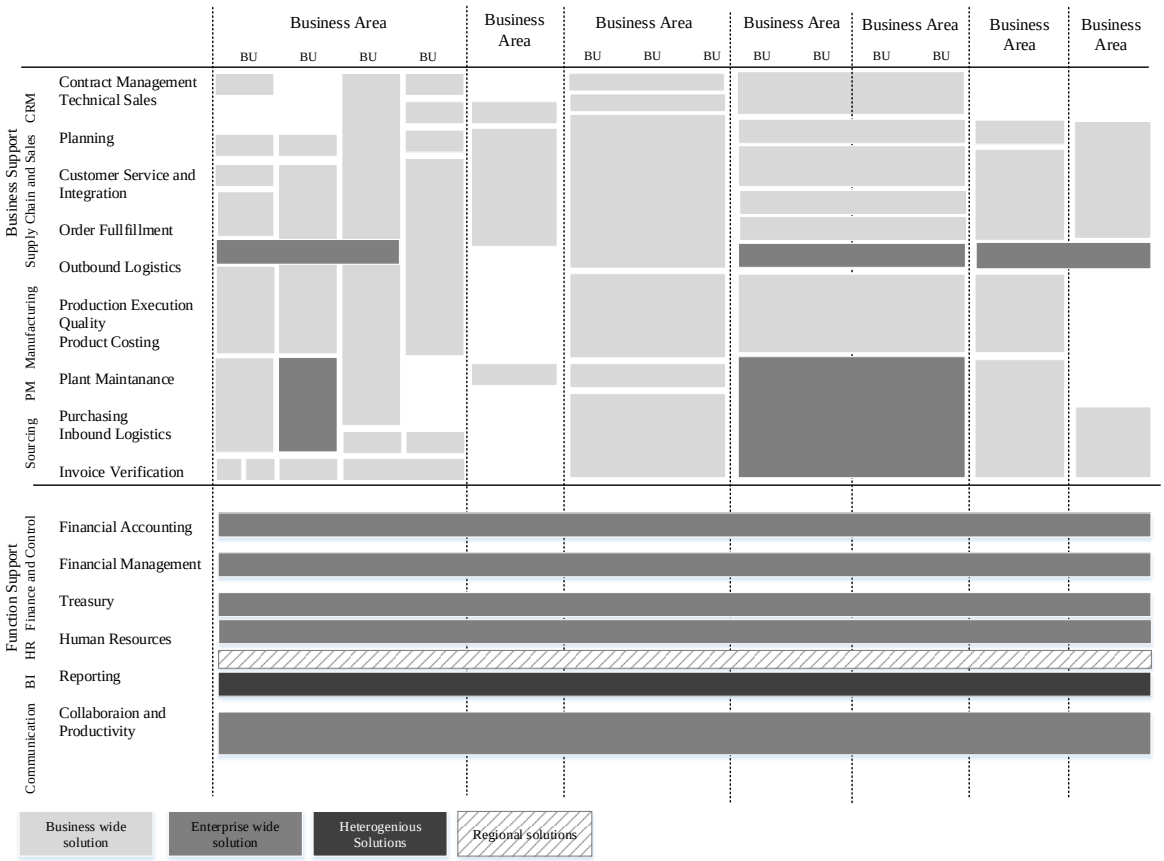

Figure 13. IT infrastructure of the case company

The third characteristic of this case are frequently changing requirements. As concepts business analytics and business intelligence are not new, but the methods and technologies are developing rapidly, which leads to changes in infrastructure. Operating in a dynamic business environment also puts pressure on companies to come up with new sources of competitive advantage. As a result companies are constantly testing and developing new ways to utilize information assets in steering of the business. The BI infrastructure needs to be flexible enough to allow the implementation of the new ideas, and also adjustment of previously implemented structures.

The combination of these three characteristics makes the case company interesting for BI research, but there surely is a multitude of companies that share these three features, or some of them, thus extending the applicability of the results to beyond the case company.

4.3 Case Company Requirements

The data collection was carried out during a period of 12 months, from the beginning of April 2016 until the end of March 2017. The data was collected by participant observation as a part of an analyst team in the case company. Observations are based on the daily tasks of the analysts, reporting development projects and by participating in meetings and workshops. In addition to the observations, documentation of current reporting practices were reviewed to understand how reporting and analysis is currently conducted. These documents include spreadsheets used by the analysts as well as documentation of IT infrastructure and data content in different systems. The following findings were made based on the collected data, which then were used to form the requirements of the case company.

4.3.1 Business Intelligence Infrastructure Requirements

By following the daily tasks of analyst it became clear quite early in the observation period that a lot of time and effort went to manual data preparation. As the data collection proceeded it also became apparent that this was limiting value creation of business analysis. Since transforming data for descriptive analysis required so much time and effort, there was no

time left for more value adding predictive, or prescriptive analysis. Based on reflection with literature two causes for the excessive amount of manual data preparation were identified: Lack of automation, and lack of reuse. A lot of analysis is done in spreadsheets using data extracted from ERP systems and other sources. In many cases most of the reporting logic could be automated and manual work reduced by better utilization of BI tools. The lack of reuse is largely connected to the lack of automation. Automating reporting logic would require standardization, which would lead to better reusability of standardized data and reporting logic. Manually managed spreadsheets are difficult to reuse for other purposes since they are usually in a non-standard format, but automated ETL scripts can be more easily reused to reduce manual work in other reporting processes than what they were originally developed for. If time consuming manual work was reduced time to insight would also be shorter, which would improve ability to respond to business requirements. Increasing automation and standardization would also improve data quality, which would increase readiness to engage in predictive and prescriptive analysis that might require large quantities of high quality data. One of the major contributors to the amount of manual work is the lack of integration between source systems. Whenever analysis is carried out that requires data from multiple sources the data needs to be manually transformed in a spreadsheet. Therefore, developing automated ETL-processes to integrate data from different systems would be especially beneficial. However, considering the complex and heterogeneous IT infrastructure of the case company, building and enterprise data warehouse could be costly and difficult, and more flexible solution would be needed.

For the most part reporting to the end users is static. Analysts prepare the data, conduct the analysis and provide the results to business users in a static format. The business users are commonly unable to explore the data themselves. This is the process for most of the reporting, but there are exceptions. The company does have BI systems for a specific subject areas that are fully integrated in to source systems and allow the business users to drill-down, pivot, and slide-and-dice the data themselves. However, these systems are costly and time consuming to develop. There is clearly a need to find ways to develop systems that provide dynamic functionalities to the business users with lower costs and better flexibility.

There were four major findings made based on the collected empirical evidence. First, BI technology could be utilized more effectively. Amount of manual work is excessive and could be reduced by proper use of BI technologies. Second, data preparation and analysis

could be more standardized. This would increase efficiency through reuse and scalability. Third, when BI technology is used the systems are expensive, slow to develop and difficult to modify. Fourth, in most cases reports are static and user engagement is minimal, which limits pervasive use of the analysis. By reviewing related literature solutions were identified that can increase the value creation of business analysis in the case company.

4.3.2. Content Requirements

Variable costs are the product of two components, price and consumption. These components, and drivers that impact them are outlined in figure 14. This so called Variable cost bridge – model was created to understand the content needs of the analysis which is being developed. The model connects the different factors that determine variable costs, and in the end, impact profitability. Market prices of the raw materials are determined by supply and demand in the global commodity markets. The market price has an impact on the purchase price of the manufacturing company, but often the prices are not the same, or do not change in the same proportion and time. The purchase price is also affected by currency rates, and the firm's own actions, such as sourcing strategy. The purchase price then forms variable costs in combination with consumption. In addition to price related drivers there are also multiple factors that impact consumption. These drivers are related to the manufacturing process, including efficiency and yield and the decisions on which products are manufactured. In the combination of all of these factors results in the EBIT impact of variable costs.

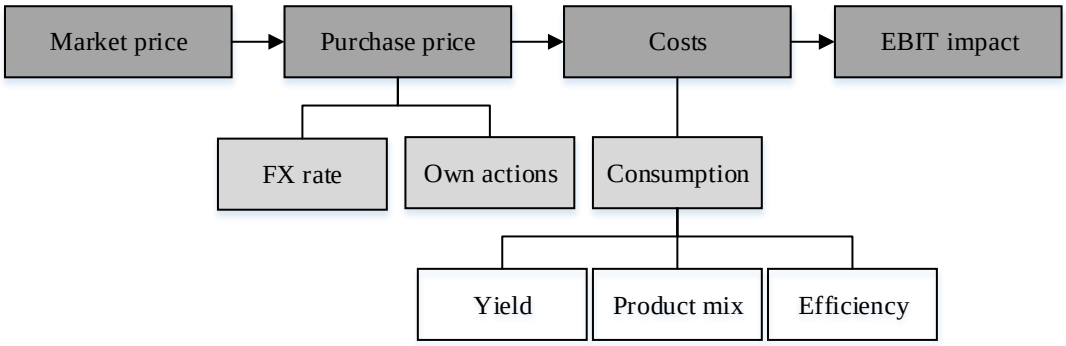

Figure 14. Illustration of the Variable Cost Bridge

The objective of the developed reporting is to provide transparency to the components of the variable cost bridge without sacrificing the agility of the system. Finding the balance between content and agility means estimating the feasibility of how much content is included and in what level of detail. The content should respond to business requirements, but not make the reporting logic too complicated so that it can be adjusted when the requirements change.

Many components of the variable cost bridge are included in different reporting solutions in different levels of the company's businesses. In the case of higher level enterprise-wide reporting all components are not of equal importance. The main priority of this variable cost analysis is to separate the market driven change in variable costs, from the impact of the company's own actions.

5. Development Process

This chapter outlines the development project that was carried out to test the chosen methods and technologies in action. The development process started with a comparison of the case company's requirements and development methods and technologies suggested by the related literature. Fist, the decisions about how the project should be managed and what technologies and tools would help to reach the objectives the case company. Second, data discovery was conducted to map the data sources, data content, and limitations. The technological selections were also verified to fit the discovered data sources and data. Last, a plan for the development was outlined.

5.2 Development Plan and Data Discovery

The main requirements of the case company regarding the BI infrastructure were agility, automation, reusability, improved data integration and more dynamic reporting. Iterative development, self-service, in-memory technology and data-blending were identified as the agile methods and technologies that would help to fulfill these requirements.

Based on empirical findings and the related literature it seemed very unlikely that developing the BI infrastructure based on traditional methods and technologies would reach the objective of this study. The case company had developed data warehousing solutions before and learned that these projects can be difficult to complete in time, and in budget. This view was also supported by literature. The fact that the variable cost report requires integration of multiple data sources supported the decision not to build a complicated data model and ETL-process. It was clear that the solution needed an agile data infrastructure. Therefore, data integration was carried out using in-memory technology that allows agile data integration. In order to limit costs and lead time of the project the integration was developed using a graph based ETL engine that allowed development to be done as self-service. Also the analysis application needed to be capable to handle frequently changing business requirements. The BI application was chosen based on self-service capabilities, so that making changes to the dashboards would be flexible and effortless. Because of the focus on agility and self-service the development work could be carried out with minimal IT

involvement. The development was conducted in iterative fashion, typical for agile development, to enable flexibility during the project. This approach also ensured that the solution delivered value to the end users.

The objective of the developed reporting is to provide transparency to the components of the variable cost bridge without sacrificing the agility of the system. Finding the balance between content and agility means estimating the feasibility of what content is included and in what level of detail. Including all possible data content at high level of detail will likely make the system unnecessarily complex without providing added benefit over less content and more flexible system.

The most important data source for the report is a financial management system used for management accounting. The system includes the data which is used in monitoring financial measures including variable costs. Therefore, this will be the primary data source used for variable cost reporting.

Multiple other data sources were also needed in order to fulfill the content requirements of the case company. Market price data was loaded from three different sources. These include two different market data vendors, as well as data that is manually gathered from sources that cannot be accessed as a database. The case company uses some specialized materials, which market prices are not easily available and therefore to gather all the needed information multiple sources were needed. Many of the materials have different market prices for different product qualities, continents, delivery terms et cetera. Identifying the correct market prices required expertise and was done in collaboration with end-users.

The own impact of the case company is estimated based on the cost impact of savings projects. The projects are monitored in a savings project managements system. Without this centralized way of monitoring the impact of cost saving actions the estimation of the case company’s own impact would be very difficult. The savings project managements system is not easy to integrate to the other data sources, but this is outweighed by the benefit of having the data as an external input, and no need to build any calculation logic.

Consumption related data can be found in multiple different sources in the case company. The sources vary on the level of detail and data content. Raw material consumption for the largest business areas can be found in a data warehouse, and consumption data for other business areas is available in each businesses ERP system. Some high level consumption

data is also available in the financial management system. Accessing these data sources is easy. However, the drivers behind consumption changes require more data to monitor. Measuring the impact of yield, product mix and efficiency on consumption could be done by including data from the case company's production systems. However, these systems are highly customized and difficult to integrate to other sources without specialized knowledge of the systems. The integration of these systems would require heavy involvement of IT and possibly consulting as well. This would increase the costs and time of the project significantly. Therefore, the consumption drivers are excluded from the content scope of the project at this point.

Most of the identified data sources are easy to connect to with the ETL-tool. The data in the sources is high quality and therefore no special attention needs to be given to cleaning of the data. The dimensions in the data can be different and some assumptions will have to be applied to the transformation process in order to integrate the data sources. The raw data of the financial management system is fairly complex and some time needed to be put in order to build understanding of the data structure. Also the raw consumption data in the data warehouse and ERP systems required understanding that had to be gained.

The actual development was made in iterations so that the process will be flexible if business requirements change during the process. Also each iteration will provide the users with valuable additions to the reporting. The iterations will be split by data source. In the first iteration the only data source that will be needed is the financial management system. The second iteration will focus on enriching the financial data with market price data. Next, savings data will be added, and in the last iteration consumption data will be included to the solution. During each iteration new content was provided for the end users. The content was tested and verified by the end-users and feedback was given. To promote reuse the development process was designed so that each iteration provided a new module to the ETL-process. These modules can be later reused in different purposes.

Documentation of the plan and business requirements was done in agile, minimalistic fashion. The goal was to minimize ceremonial bureaucracy. A rough plan of the final application and data integration plan for each iteration was documented, but no detailed plan was produced before the start of development work since the details were assumed to change during the project. This documentation practice allowed keeping the documentation up to

date without considerable effort, and made it easier to focus on creating business value. A high level chart of the planned architecture is presented in figure 15.

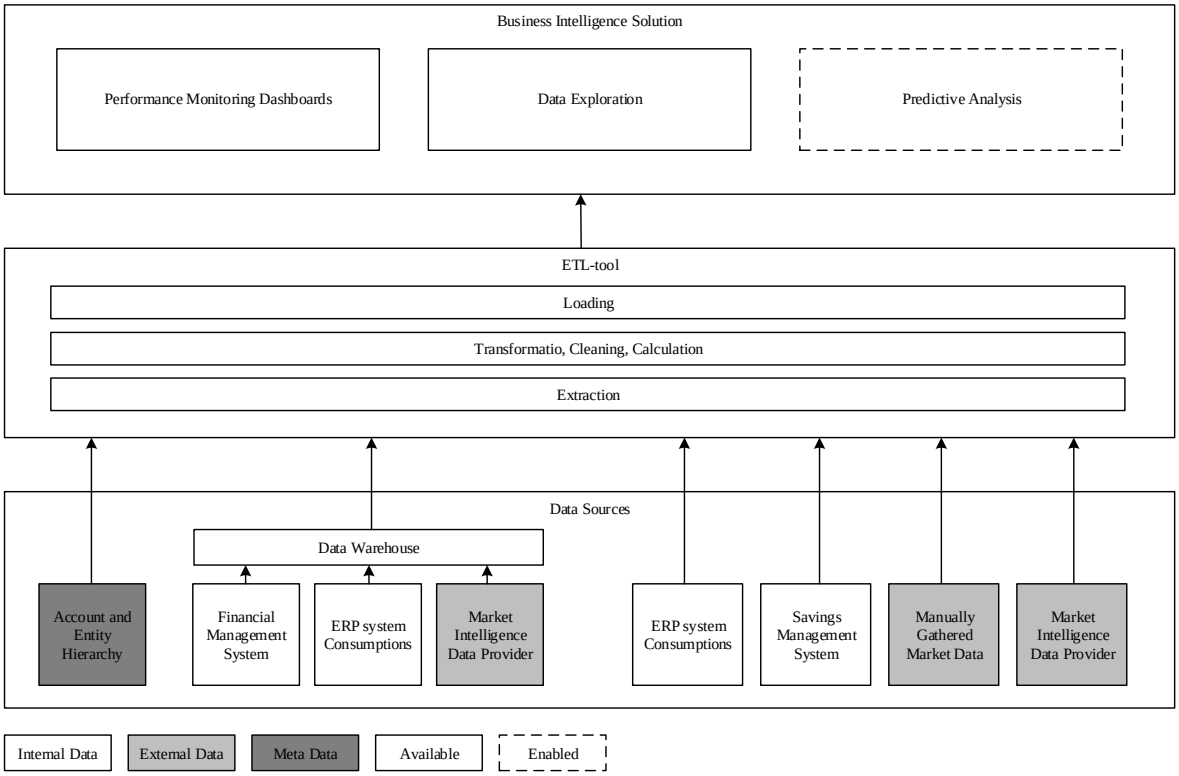

Figure 15. Planned BI architecture

5.4. Development Iterations

In this section the work done during each development iteration is presented. The purpose of these chapters is to provide the reader information which can be used to assess the project in which the methods and technologies where tested in, what type of challenges where included and what kind of analytical logic was implemented. By using this information the reader is able to have a clear understanding of the requirements of conducting group level business analysis in the case environment and what challenges do the used methods and technologies resolve.

5.4.1 Group Variable Costs

In the beginning of the development process focus was on implementing the automated data preparation process and dynamic reporting platform. In this phase the only data source was the financial management system which data will be enriched in the subsequent iterations. The goal of the first iteration was to create value by providing the end user with a more advanced reporting platform and by releasing time from the analysts by implementing an automated reporting process.

The raw data from the financial management system had not been used before in an external reporting platform and majority of the time and effort in the first iteration went to building understanding of the data and planning what kind of steps were needed to prepare it to the reporting layer. To keep the reporting layer as simple as possible most of the metrics needed in the reporting were calculated before the data was loaded into the application. There were no issues in the data extraction from the financial management systems database. The data was extracted to the ETL-tool by an SQL query and the actual data preparation was done in-memory of the tool. The raw data in the financial management system was over 3,7 million records. Including three years of historical data. This is one of the reasons why the analysis would be rather difficult to do in a spreadsheet, and more advanced BI system is required. The scope of the first iteration did not require all of the data records and therefore the data was filtered down to only the required entities and accounts. Rigorous testing was made thorough out the development process to make sure filters applied in the ETL-tool did not exclude any needed data, or include anything unnecessary that would result in incorrect outcomes in the reporting layer when the data is aggregated.

In the case company variable costs are commonly measured as costs per production volume. Both variable costs and production volumes (delivery volume for delivery costs) are included in the data, which allows reporting variable costs per volume. In addition to the changes and development in variable cost, there are two metrics calculated before the reporting layer.

Unit Cost Change measures the change in variable costs versus the comparison period without taking changes in production volume into account. It is calculated as:

$$\text{Unit Cost Change} = \left(\frac{c_n}{p_n} - \frac{c_{n-1}}{p_{n-1}}\right) \times p_{n-1}$$

where, c is variable cost and p is production volume.

Volume Impact measures the impact of production volume change versus the comparison period. It is calculated as:

$$Volume\ Impact = (p_n - p_{n-1}) \times \frac{c_{n-1}}{p_{n-1}}$$

where, c is variable cost and p is production volume.

The case company is a group of several different business areas which are further divided into business units and production units. The entity hierarchy is not consistent, but some units can be split to product categories whereas some to production lines. This makes the entity and account hierarchies of the financial data rather complicated. There has also been changes in the organizational structure that need to be accounted for as well as new and closed production units. The business units also do intercompany trade which needs to be eliminated when the costs are aggregated to higher level of the hierarchy.

5.4.2 Market Data

After the initial implementation of the BI infrastructure and reporting process, the second increment is focused on enriching the financial data with market related information. The internal data is integrated with external market data coming from three new sources. The data extracted from the sources includes market prices for raw materials and feedstocks, as well as currency exchange rates. Also currency exchange rate cost impact, which is calculated in the financial management system, is now included in the data that is extracted from that database. These datasets will provide the company with better visibility on what kind of impact is the supply market having on their variable costs. The information can be used to monitor the company's sourcing performance, as well as make strategic decisions regarding the future. However, the data integration is not very straightforward. There are significant limitations that needed to be considered when integrating internal and high level financial data, with external and detailed market data.

Two of the market data sources are data bases which data is purchased from external data vendors and the third source is a flat file which includes price data that is not available in

any database. Data provided by the first market data vendor is currently loaded to the company's market intelligence database as a monthly process. Accessing this information required some assistance to from the company's IT department, to identify the correct tables and fields in the database. Actual development work was rather effortless and was done completely as self-service. The market data from the market intelligence database also includes currency exchange rates. The data provided by the second market data vendor is not loaded to the company's market intelligence database, but can be accessed as a web service in the vendors database. Accessing and developing the data integration to the second market data source was done as self-service from start to finish.

The supply market is one of the major drivers impacting the company's variable costs and therefore market impact is an important metric. Estimating the market impact however is not easy. A completely accurate estimate for market impact is impossible to produce, because there are so many factors that would have to be considered. The chosen calculation is easy to understand and implement. It provides a high level estimate of the impact, and a more advanced calculation can be implemented later when there is more understanding of how the impact could be measured more accurately. At this point market impact is measured as:

$$Market\ Impact = \frac{m_n - m_{n-1}}{m_{n-1}} \times (c_n - c_{n-1})$$

where, c is variable cost and m is production volume.

Changes in currency exchange can also have a significant impact on variable costs. Currency Impact measures the impact of changes in currency exchange rates versus the comparison period. This is the only metric used that is calculated in the financial management system and is already included in the data that is extracted from the database. Currency impact is calculated as:

$$Currency\ Impact = \frac{f_n - f_{n-1}}{f_{n-1}} \times (c_n - c_{n-1})$$

where, c is variable cost and f is foreign exchange rate.

For most cost categories the foreign exchange rate is the exchange rate between the local currency where the unit is located and the case company's home currency. One of the major cost categories has its own calculation where the local currency is replaced by US Dollar since that is the currency which is used to measure the value of this raw material globally.

Integrating the financial data and market data had two major limitations. Firstly, the case company purchases thousands of different materials. It is not possible to estimate the impact for all, and many of them do not even have a market price. The market prices can also be very different based on quality of the material or geographical area. Secondly, the financial data, which is used primarily for accounting purposes, is low granularity data including cost groups, and no information on individual materials. Therefore, it is not possible to find matching market prices for each cost category. Multiple prices are required per category, and knowing which ones requires expertise.

Solving these two issues required close co-operation with the businesses and sourcing managers who are familiar with the raw materials and cost structures of the company. The market prices that are equivalent to what the company's raw materials where chosen, and combined with the corresponding cost categories. The solution was also developed to allow changing market prices as the requirements change. This flexibility will be needed if the market data vendors are changed, or the company's cost structure changes significantly.

5.4.3 Own Impact

This iteration provided visibility on the impact of cost saving actions the case company is taking. The savings data was extracted from a savings project management (SPM) system, which is used by the case company to monitor the savings projects. Savings combined with the market impact measures implemented in the previous increment, delivered transparency on the most important variable cost drivers. However, the integration of this data added some manual work to the ETL-process.

The SPM system which was used as a source for the own impact is a SaaS (Software as a Service) solution and could not be directly connected with the ETL-tool. Thus, the data needs to be extracted manually from the system to a flat file which the ETL-tool is reading. This was not completely in line with the requirements, because of the required manual work. However, the case company has accepted this approach for now, bending possible changes to the savings management system by the service provider that would allow a direct data connection.

The case company is engaged in various types of savings projects, which are input manually to the system by the employees who are responsible for them. All tough the manual input process is prone to human error, the quality of the data is high. Some changes in the case company's business structure are not implemented in the historical data of the SPM system and there needed to be some reallocation of the historical savings according to the current business structure.

The data content has SPM system been designed to be comparable with the financial management systems data and the two datasets share the same dimensions which makes the integration work easier. There are two metrics included in the SPM data, which are realized savings and potential savings. Realized savings are savings from implemented savings projects that have already been realized. Potential savings are estimations of savings from already implemented or planned projects that are expected to be realized in the future. Because of the variation in the types of savings projects that can be undertaken and the way that savings from these projects can be calculated, there is no calculation logic build in to the SPM system. Realized and potential savings are calculated outside of the system by the employees who are responsible for the project and then input to the system. The system then does monthly cost impact allocations of the savings based on the time period specified by the project owner.

Savings programs are managed in the same manner by all of the case company's businesses. However, one of the business areas was not using the SPM system, but has their own manual savings monitoring practice. This business area will start using the SPM system in the future, which means that there will need to be some changes made to the data preparation process. Fortunately, the agile technologies make the process easy to adjust.

5.4.4 Material Consumption

In the last one of the planned development iterations the objective was to provide visibility on consumption changes. Consumption is one of the key elements of the variable cost bridge model and an important part of the analysis. Unfortunately providing this information for all materials used group-wide would be rather difficult, and therefore some limitations where implemented. By utilizing this information the end user can analyse how consumption

changes have impacted variable costs. However, the drivers behind the consumption change will not be visible. Even so, consumption information is valuable information for variable cost analysis.

Material consumption numbers were not directly available in any system or database, but were calculated from material movements that have been recorded to the case company's ERP systems. Material movement data for the largest business areas is loaded in to a data warehouse. Some of the business areas are so small that changes in their material consumption will not have a significant impact on variable costs from group perspective, therefore only one business area specific ERP system was integrated in addition to the data warehouse data.

The material movement data includes multiple types of movements and many different locations and cost centers. Only specific types of movements and cost centers are needed to calculate the raw material consumption in production processes. Finding the right movement types and cost centers from the raw data required some support from specialists. After unnecessary material movements where filtered from the data calculating consumption only required subtracting one movement type from another for each cost center.

The case company uses multiple types of materials, many of which are measured in different units. This has to be considered when the BI application layer is developed. Aggregating different units of measures is not wise. Also visualizing changes in different sized units of measures can be misleading. Therefore, consumption changes were measured on per produced output basis. Reporting consumption per produced unit of output gives the user a better understanding of how different materials have been consumed in production. By using this information the user can make assumptions on the drivers of consumption change. If the weights of the consumed materials has changed, it is likely that product mix has been changed since production of different products consume materials in different proportions. If the consumption volume per produced unit has generally changed it is likely that yield or efficiency of the production process has changed.

5.5 Overview of the BI Application

Dashboards in the BI application where designed to support the performance monitoring process. The application provides both high level information and more detailed information in different dashboards. Therefore, the user is able to update high level situation awareness efficiently, but can also view the information in more detail when needed.

The first dashboard provides an overview of variable cost related metrics on group level. The objective of the group overview dashboard is to update the high level situation to the user. Since some of the metrics cannot be aggregated to group level this dashboard will not provide all metrics on group level, but some metrics are reported on business area level which also allows the user to compare performance in each business area. Included in the bottom of the overview dashboard are the main cost drivers of the variable cost bridge. These are currency impact, market impact, savings and volume impact. The bottom part of the overview dashboards presents the impact of variable cost drivers on specific cost components, as well as total costs for the components, change from previous year and outlook for the whole year.

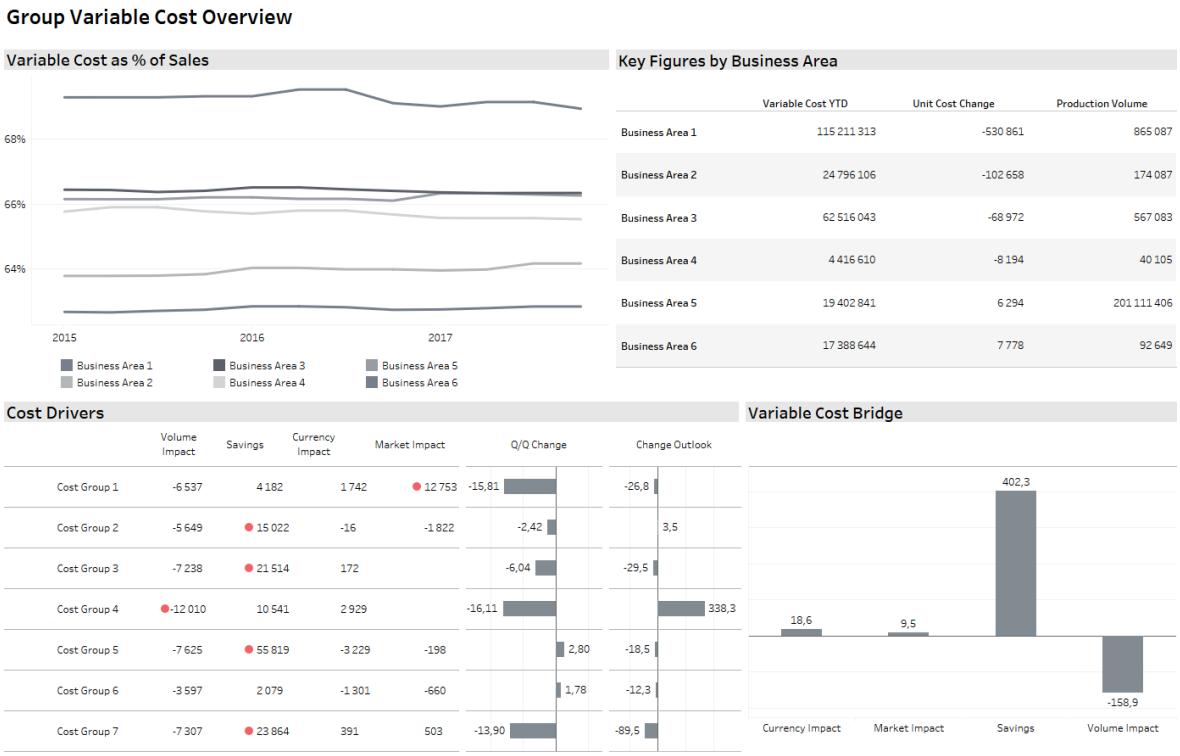

Figure 16. Group Variable Cost Overview dashboard

The view can be tailored by selecting a specific entity from the top of the dashboard. The top part presents selected high level KPI's including their trend and comparison to the previous year, as well as outlook for the whole year. The main cost drivers of the variable cost bridge are included also in the business overview dashboard, and again also the part of the change in variable costs which is not explained by these figures is provided. The bottom part of the business overview dashboards presents the impact of variable cost drivers on cost components on more detailed level than the group overview, as well as total costs for the components, change from previous year and outlook for the whole year.

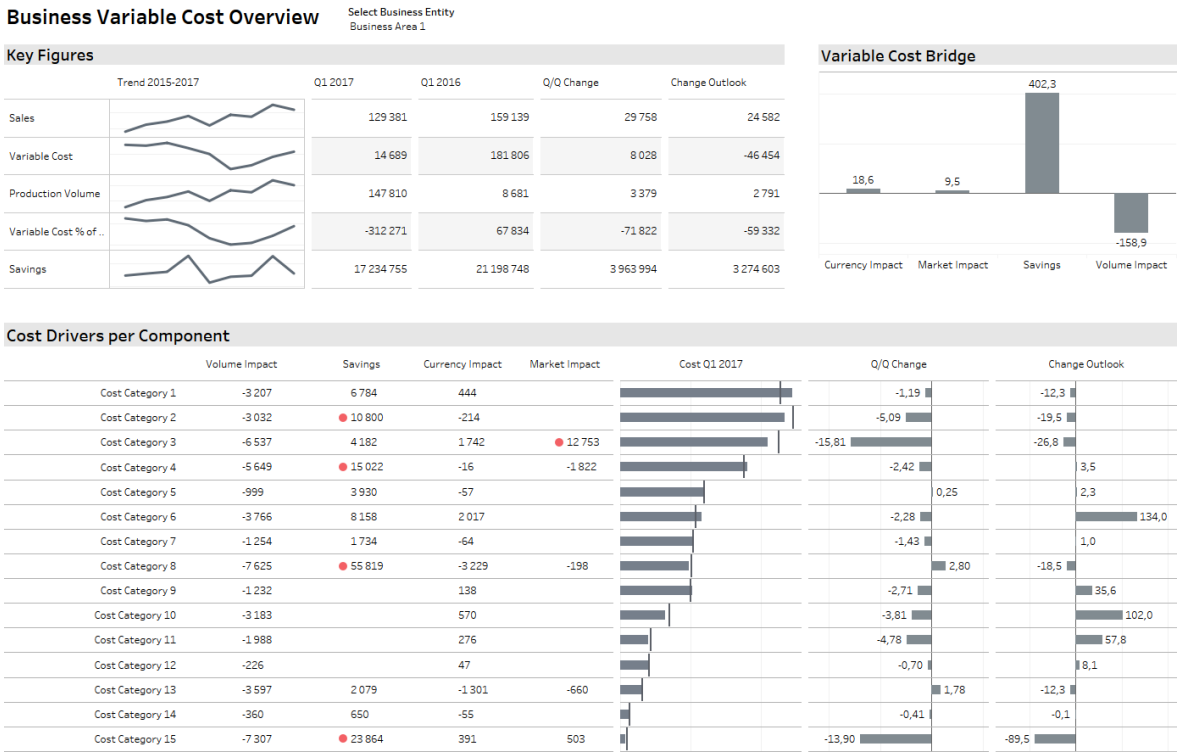

Figure 17. Business Overview dashboard

Whereas the first two dashboards are designed for monitoring performance on a higher level, the purpose of the third dashboard is to facilitate more detailed analysis. If the user identifies specific business or cost component that needs further attention, due to a large change for example, the variable cost driver dashboard can be used to analyse in greater detail what is causing the change. The cost driver dashboard presents market price trends, currency exchange rates, production volumes, savings, and consumptions, which are the drivers of variable costs. The dashboard focuses on visual presentation of data in order to provide the user with the ability to visually analyse trends and patterns in the data. The user can also explore the data by filtering the dashboard for a specific business, and cost category.

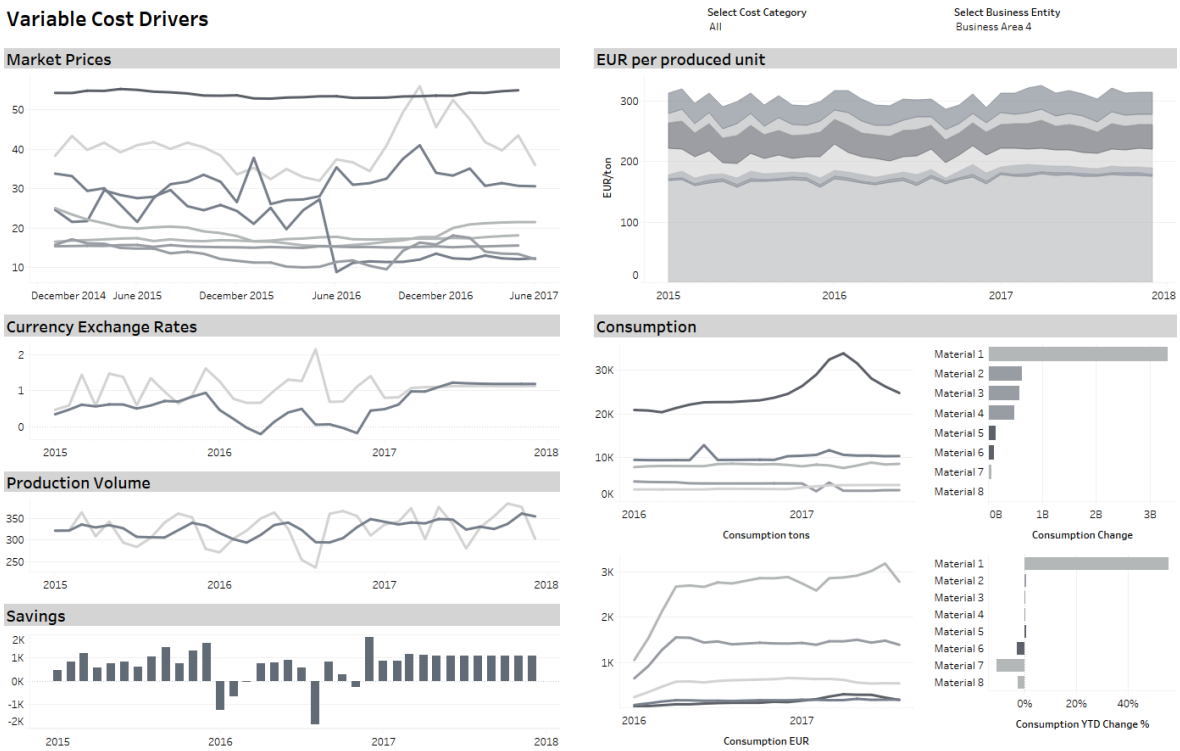

Figure 18. Variable Cost Driver dashboard

In case additional analysis is required the data used in the application is available for analysts to use. These functionalities are not made available for all end users, but only for ones with data analysis capabilities. Therefore, if the dashboards do not provide the information needed, an analyst is able to carry out further analysis, build own visualizations, or use statistical models with the data. The fact that the data can be easily used by analysts for other purposes as well is one of the main benefits of developing the solution as self-service. The analysis that is currently possible in the solution is simple descriptive type of analysis, but if usable statistical models are identified they can be easily developed in to the application. Thus, the solution enables higher value creation with predictive models.

6. Discussion

The case company was looking to build a platform that would enable higher value creation with business analysis. Previous literature suggested that focusing on timely delivery of insights, pervasive use and predictive analysis would help create higher value. A BI system was developed based on the requirements of the case company using one of the most difficult topics as an example. The main requirements of the case company regarding the BI system were automation, reusability, improved data integration and more dynamic reporting. These were identified based on observations made in the case company and review of previous studies. The development work was to be done following agile principles. Iteration, self-service, in-memory technology and data-blending were identified as the agile methods and technologies that would help to reach the case company's objective.

Many corporations that have corporate information applications still lack the information needed for decision making, because the information in the applications is unintegrated, there is no historical data available, or data is not easy to get to even in the applications (Rasmussen et al. 2002). As a multi-business enterprise with a complex IT infrastructure also the case company has issues with using their data efficiently. There are some implementations of enterprise wide solutions, but these have been difficult and costly to develop and maintain. It is not a surprise that integrated analytical databases are not that common in the case company. In many cases historical data is not available or data quality is not satisfactory for analytical purposes. For these reasons business analytics requires a lot of manual data preparation at the moment. This particular case of group variable cost analysis required the integration of multiple data sources. Therefore, it was decided that an agile business intelligence system would be the best way to go forward in this case.

Previous literature suggests that all types of business analytics do not create equal levels of value. The more advanced the analysis is the more value it potentially creates. Obviously predictive analysis requires more capabilities, better data and more advanced tools. In the past the case company has not focused on implementing predictive analytics, but has now the will to move to that direction. Therefore, the company needed a platform to enable predictive analysis. Even though statistical models are not included at this stage, the implemented platform supports advanced statistical analysis. Therefore, the case company can implement predictive analysis in the same platform where monitoring and exploration

takes place. The ability to have an agile business intelligence platform that can be used for multiple different analytical purposes increases the pervasiveness of the analysis and as a result increases value creation. (Wixom et al. 2013)

Automating as many routine processes as possible is the only way to be truly agile. Hence, truly agile teams seek to automate everything that is done more than once. (Collier 2011) Automation is one of the key factors in maximizing the value creation of business analytics through increased time to insight. (Wixom et al. 2013) The longer it takes to get from data to insight that can be used in decision making the lower the created value is (Hackathron 2004). Therefore, automation was one of the main considerations for the case company. Based on the observations in the case company a perception was formed that when routine tasks would be automated the case company's analysts would also have more time to focus on more high value adding analysis.

Another limitation that was observed in the case company was that most analyses had completely separate data preparation processes. Wixom et al. (2013) promote reuse to help maximize value creation. Therefore, the new business intelligence system was developed in a way that would allow parts of the data preparation process to be used for other analyses as well. This was found to be a highly useful development practice, since some parts of the process where implemented for other purposes already during this project.

In addition to the perceived benefits of creating an agile platform for business analytics this study also found agile development methods suitable for the case context. As many previous studies have also argued an iterative development methodology was found highly suitable for BI development. Like many other companies the case company has recognized that there is a need to find alternative ways of delivering BI solutions (Imhoff and White 2011). Since agile is a value-driven approach, which highest priority is early and continuous delivery it was decided that the development work would be done in a flexible fashion, with a low level of bureaucracy (Collier 2011). A flexible style of development did ensure that focus stayed in the delivery of valuable features. However, it was noticed that a certain level of bureaucracy is a benefit. Documentation of the development work was especially helpful when moving from one iteration to another. Nevertheless, focus delivering business value instead of elegant data models and efficient ETL-scripts was perceived as a positive change compared to previous experiences in BI system development.

Previous literature argued that the agile approach would deliver measurable results that are visible for the client, which would improve the clients ability to take ownership of the solution, ease the transfer of knowledge and make change management more efficient (Yeoh and Koronios 2010). Agile methods seek for deeper collaboration between the development team and its stakeholders, and also acknowledges that the responsibility for success is shared (Collier 2011). In this case it became evident that close collaboration between the developer and business users benefited both sides. The developer was able to better understand what features created value for the business users, and the business users understood how the solution worked, which made it easier for them to take ownership of the solution and make confident decisions based on the information it provided.

Based on previous literature one of the enablers of more efficient and more agile business intelligence environment is to enable users to become more independent, self-reliant and less dependent on the help of the IT department. Therefore, the development of the solution was done as much independent from IT as possible. Based on earlier findings the development project focused on facilitating an end user enabling business intelligence environment, ease of accessing data, ease of using the data, and flexibility of adjusting the BI solution (Imhoff and White 2011). It became clear that implementing the right self-service enabling technologies it was possible to achieve a high degree of independence from the IT department. However, as have been found in previous research not only technological investment creates value from data, but also allocation of the correct capabilities is required (Sharma et al. 2014). Especially in the complex IT environment of the case company understanding of the data and the source systems was vital.

It has been discovered that application of in-memory technology can have a significant impact on how IT systems are used (Winter et al. 2011). It can turn static decision making into dynamic iterative analysis. This is the type of change that also the case company has been looking to make into its decision making process. Agile ETL-tools are able to acquire data directly from a wide variety source systems allowing business analysts far quicker access to data by eliminating the need to build an enterprise wide data warehouse. In this case data integration was carried out using data blending in an in-memory tool. This approach proved to be highly useful to facilitate the iterative development process. BI infrastructures are often rigid, inflexible and unsuitable for readjustments (Knabke and Olbrich 2011). In order for agile business intelligence systems to be created, the whole

architecture must become more agile. Using an agile ETL-tool made the integration of heterogeneous IT systems more straight forward allowing the data to be used more efficiently.

7. Conclusions

The objective of this study was to develop a business intelligence system for the case company that responds to the requirements of conducting high value creating business analysis. The research process started with an observation period during which the current practices and systems in the case company were assessed. The purpose of the observation period was to have a thorough understanding of what is limiting value creation of business analysis in the case company. The assessment was made based on what prior research had identified as value maximizing factors. Methods and technologies suggested by literature to facilitate high value creation were chosen and those methods and technologies were put to a test using one of the most difficult topics in the case company.

Previous literature suggested that focusing on timely delivery of insights, pervasive use and predictive analysis would help create higher value. The main limitations in the case company were lack of automation and reuse, poor data integration and static reporting. Iteration, self-service, in-memory technology and data-blending were identified as the agile methods and technologies that would provide improved time to insight, and a dynamic reporting platform would make using the system more pervasive. The platform would also need to be able to support predictive analysis. Implementing these methods and technologies provided significant improvement compared to previous experiences of the case company. Even so, the implementation of the system still required specialized expertise and thorough understanding of the data. Also the significant time used to study the situation in the case company helped to identify the solutions to implement.

There are no case studies in the previous literature that would focus on agile BI methodology and technology which would include detailed descriptions of the development work. This study provides practitioners with information that can be used to assess the project in which the methods and technologies where tested in, what type of challenges where faced and what kind of analytical logic was implemented. By using this information the reader is able to have a clear understanding of the requirements of conducting group level business analysis in the case context and what challenges do the used methods and technologies resolve. Practitioners can use this study as a reference in the future when planning the implementation of analytical platforms. It has to be remembered that as a case study many of the details in

this study are highly case related and the findings are not valid in all situations. However, similar case contexts can be found where the main conclusions of this study do apply.

Business analytics and business intelligence related literature is ample. As business is conducted in increasingly dynamic environment also agile methods have gained more attention in the recent years. In order to help implement the knowledge of agile methods and technologies academic literature should focus more on practitioner-oriented research. There is a need among companies to help understand the challenges in using agile project management methods and technical practices. Research should shift towards more practical approach. Not only regarding agile, but especially regarding upcoming technologies such as data lake and machine learning that are spreading to the business community, but are not yet generally well understood.

References

Acito, F., & Khatri, V. (2014). Business analytics: Why now and what next?. Business Horizons, 57(5), 565-570.

Agrawal, R., Gupta, A., & Sarawagi, S. (1997). Modeling multidimensional databases. Data Engineering, 1997. Proceedings. 13th International Conference on (pp. 232-243). IEEE.

Ambler, S. W. (2003). Agile model driven development is good enough. IEEE Software, 20(5), 71-7

Ambler, S. W. (2007). Test-Driven Development of Relational Databases. IEEE Software, 24(3), 37.

Ambler, S. (2012). Agile database techniques: Effective strategies for the agile software developer. John Wiley & Sons.

Alagić, S. (2015). Object-Oriented Technology. Springer.

Alteryx. (2015). The Definitive Guide to Data Blending. [White Paper].

Anderson-Lehman, R., Watson, H. J., Wixom, B. H., & Hoffer, J. A. (2004). Continental Airlines flies high with real-time business intelligence. MIS Quarterly Executive, 3(4), 163-176.

Ariav, G., & Ginzberg, M. J. (1985). DSS design: a systemic view of decision support. Communications of the ACM, 28(10), 1045-1052.

Baars, H., & Hütter, H. (2015, January). A framework for identifying and selecting measures to enhance BI-agility. In System Sciences (HICSS), 2015 48th Hawaii International Conference on (pp. 4712-4721). IEEE.

Baars, H., & Kemper, H. G. (2008). Management support with structured and unstructured data—an integrated business intelligence framework. Information Systems Management, 25(2), 132-148.

Baars, H., & Zimmer, M. (2013). A Classification For Business Intelligence Agility Indicators. In ECIS (p. 163).

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. The qualitative report, 13(4), 544-559.

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., et al. (2001). Manifesto for agile software development.. Available on: http:// agilemanifesto.org/

Bruni, M. (2011). 5 Steps To Agile BI. Information Week, (1303), 33-34.

Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. ACM Sigmod record, 26(1), 65-74.

Chaki, S. (2015). Enterprise Information Management in Practice. Apress.

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

Codd, E. F., Codd, S. B., & Salley, C. T. (1993). Providing OLAP (on-line analytical processing) to user-analysts: An IT mandate. Codd and Date, 32.

Collier, K. (2011). Agile analytics: A value-driven approach to business intelligence and data warehousing. Addison-Wesley.

Cook, T. S., & Nagy, P. (2014). Business intelligence for the radiologist: making your data work for you. Journal of the American College of Radiology, 11(12), 1238-1240.

Cooper, A. (2012). What is analytics? Definition and essential characteristics. CETIS Analytics Series, 1(5), 1-10.

Davenport, T. H., & Harris, J. G. (2007). Competing on analytics: The new science of winning. Harvard Business Press.

Davenport, T.H., Harris, J.H., and Morison, R. (2010). Analytics at Work: Smarter Decisions, Better Results. Boston, MA. Harvard Business Press.

Davenport, T. H., De Long, D. W., & Beers, M. C. (1998). Successful knowledge management projects. MIT Sloan Management Review, 39(2), 43.

Davenport, T. H. , Harris, J. G. , de Long, D. W. , & Jacobson, A. L. (2001). Data to knowledge to results. California Management Review, 43 (2), 117–138 .

Davenport, T.H., (2013). Analytics 3.0. [Online document]. Available: https://hbr.org/2013/12/analytics-30 Accessed: 21.12.2016.

Davenport, T. H. (2014). Keep up with your quants. Harvard Business Review, 91(7/8).

Delen, D, & Demirkan, H. (2013a). Data, information and analytics as services. Decision Support Systems, 55(1), 359-363.

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

DeSarra, P. (2012). BI Dashboards the Agile Way. Business Intelligence Journal, 17(4), 14.

Deshpande, K., & Desai, B. (2015) Model for Assessment of Agile Methodology for Implementing Data Warehouse Projects.

Eckerson, W. (2003). Smart companies in the 21st century: The secrets of creating successful business intelligence solutions. TDWI Report Series, 7.

Eisenhardt, K. M. (1989). Building theories from case study research. Academy of management review, 14(4), ss. 532-550.

Elbashir, M. Z., Collier, P. A., & Sutton, S. G. (2011). The role of organizational absorptive capacity in strategic use of business intelligence to support integrated management control systems. The Accounting Review, 86(1), 155-184.

Emblemsvåg, J. (2005). Business analytics: getting behind the numbers. International Journal of Productivity and Performance Management, 54(1), 47-58.

Endsley, M. R., & Garland, D. J. (2000). Theoretical underpinnings of situation awareness: A critical review. Situation awareness analysis and measurement, 3-32.

Evans, J. R., & Lindner, C. H. (2012). Business analytics: the next frontier for decision sciences. Decision Line, 43(2), 4-6.

Evelson, B. (2011). The Forrester Wave™: Self-Service Business Intelligence Platforms, Q2 2012. Intelligence.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine, 17(3), 37.

Few, S. (2013). Information Dashboard Design: Displaying data for at-a-glance monitoring. Burlingame, CA: Analytics Press.

Foster, E. C., & Godbole, S. (2016). Database systems: a pragmatic approach. Apress.

Garcia-Molina, H., & Salem, K. (1992). Main memory database systems: An overview. IEEE Transactions on knowledge and data engineering, 4(6), 509-516.

Gibson, M., Arnott, D., Jagielska, I., & Melbourne, A. (2004). Evaluating the intangible benefits of business intelligence: Review & research agenda. In Proceedings of the 2004 IFIP International Conference on Decision Support Systems (DSS2004): Decision Support in an Uncertain and Complex World (pp. 295-305). Prato, Italy.

Goodhue, D. L., Wixom B & Watson, H. J. (2002) Realizing business benefits through CRM: hitting the right target the right way. MISQ Executive 1(2), 79–94.

Gorry, G. A., & Morton, M. S. S. (1971). A Framework for Management Information Systems. Sloan Management Review, 13(1), 55.

Hackathorn, R. (2004). Real-time to real-value. Information Management, 14(1), pp. 24-28.

Harris, J., Craig, E., & Egan, H. (2010). How successful organizations strategically manage their analytic talent. Strategy & Leadership, 38(3), 15-22.

Hovis, G. (2002). Stop searching for information–monitor it with dashboard technology. 2013-12-14]. http://www. information-management. com/infodirect/20020208/4681-1. html.

Hughes, R. (2008). Agile Data Warehousing: Delivering world-class business intelligence systems using Scrum and XP. IUniverse.

Hughes, R. (2012). Agile data warehousing project management: business intelligence systems using Scrum. Newnes.

IBM. (2013). Descriptive, predictive, prescriptive: Transforming asset and facilities management with analytics. Thought Leadership White Paper.

Imhoff, C., & White, C. (2011). Self-service business intelligence. Empowering Users to Generate Insights, TDWI Best practices report, TWDI, Renton, WA.

Jacobsohn, M., Delurey, M., & Principal, E. (2014). How the data lake works. Booz Allen White Paper.

Jagadish, H. V., Lakshmanan, L. V., & Srivastava, D. (1999). Hierarchical or relational? A case for a modern hierarchical data model. In Knowledge and Data Engineering Exchange, 1999.(KDEX'99) Proceedings. 1999 Workshop on (pp. 3-10). IEEE.

Jank, W. (2011). Business Analytics for Managers. Springer New York.

Joia, L.A. (2000), "Measuring intangible corporate assets: linking business strategy with intellectual capital", Journal of Intellectual Capital, Vol. 1 No. 1, pp. 68-84.

Jörg, T., & Dessloch, S. (2009). Near real-time data warehousing using state-of-the-art ETL tools. In International Workshop on Business Intelligence for the Real-Time Enterprise (pp. 100-117). Springer Berlin Heidelberg.

Kankanhalli, A, Lee, O. K., & Lim, K, (2011) Knowledge reuse through electronic repositories: a study in the context of customer service support. Information and Management 48(2–3), 106–113.

Kimball, R., & Ross, M. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling. John Wiley & Sons.

Kiron, D., Prentice, P. K., & Ferguson, R. B. (2014). The analytics mandate. MIT Sloan management review, 55(4), 1.

Knabke, T., & Olbrich, S. (2011, December). Towards agile BI: applying in-memory technology to data warehouse architectures. In IMDM (pp. 101-114).

Kocakoç, I. D., & Erdem, S. (2010). Business intelligence applications in retail business: OLAP, data mining & reporting services. Journal of Information & Knowledge Management, 9(02), 171-181.

Kock, N., & McQueen, R. (1998). Knowledge and information communication in organizations: an analysis of core, support and improvement processes. Knowledge and Process Management, 5(1), 29-40.

Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). Emerging trends in business analytics. Communications of the ACM, 45(8), 45-48.

Kotu, V., & Deshpande, B. (2015). Predictive analytics and data mining: concepts and practice with rapidminer. Morgan Kaufmann.

Krawatzeck, R., Dinter, B., & Thi, D. A. P. (2015, January). How to make business intelligence agile: The Agile BI actions catalog. In System Sciences (HICSS), 2015 48th Hawaii International Conference on (pp. 4762-4771). IEEE.

Larson, D. (2009). BI principles for agile development: Keeping focused. Business Intelligence Journal, 14(4), 36-41.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. International Journal of Information Management, 36(5), 700-710.

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

Lavastorm. (2014). Why Most Big Data Projects Fail: Learning from Common Mistakes to Transform Big Data into Insights. [White Paper].

Lavastorm. (2016). Breaking Through Your Data Bottleneck With Agile ETL. [Webinar] Available: http://www.lavastorm.com/resources/webinar/. Accessed 7.12.2016.

Markus, L. M., & Tanis, C, (2000) The enterprise system experience – from adoption to success. In Framing the Domains of IT Management: Projecting the Future through the Past (ZMUD RW, Ed), pp 173–207, Pinnaflex, Cincinnati, OH.

Markus, M. L., (2004) Technochange management: using IT to drive organizational change. Journal of Information Technology 19(1), 3–19.

McNiff, J., & Whitehead, J. (2000). Action research in organisations. Psychology Press.

Muntean, M., & Surcel, T. (2013). Agile BI-the future of BI. Informatica Economica, 17(3), 114.

Morton, K., Bunker, R., Mackinlay, J., Morton, R., & Stolte, C. (2012). Dynamic workload driven data integration in tableau. 2012 ACM SIGMOD International Conference on Management of Data (pp. 807-816).

Olszak, C. M., & Ziemba, E. (2003). Business Intelligence as a Key to Management of an Enterprise. Proceedings of Informing Science and IT Education Conference (pp. 855-863).

Olszak, C. M., & Ziemba, E. (2006). Business intelligence systems in the holistic infrastructure development supporting decision-making in organisations. Interdisciplinary Journal of Information, Knowledge, and Management, 1(1), 47-57.

Ong, I. L., Siew, P. H., & Wong, S. F. (2011). A Five-Layered Business Intelligence Architecture. Communications of the IBIMA.

Pape, T. (2016). Prioritising data items for business analytics: Framework and application to human resources. European Journal of Operational Research, 252(2), 687-698.

Pedersen, T. B., Jensen, C. S., & Dyreson, C. E. (1999). Extending Practical Pre-Aggregation in On-Line Analytical Processing.

Pedersen, T. B., & Jensen, C. S. (2001). Multidimensional database technology. Computer, 34(12), 40-46.

Petrović, M., Vučković, M., Turajlić, N., Babarogić, S., Aničić, N., & Marjanović, Z. (2016) Automating ETL processes using the domain-specific modeling approach. Information Systems and e-Business Management, 1-36.

Ponniah, P. (2003). Database design and development: an essential guide for IT professionals. IEEE Press.

Power, D. J. (2001). Supporting decision-makers: An expanded framework. e-Proceedings Informing Science Conference, Krakow, Poland (pp. 431-436).

Rainardi, V. (2008). Building a data warehouse: with examples in SQL Server. John Wiley & Sons.

Rasmussen, N. H., Goldy, P. S., & Solli, P. O. (2002). Financial business intelligence: trends, technology, software selection, and implementation. John Wiley & Sons.

Reinschmidt, J., & Francoise, A. (2000). Business intelligence certification guide. IBM International Technical Support Organisation.

Robson, C. (2002) Real World Research (second edition). Oxford. Blackwell.

Rostek, K. (2010). Data Analytical Processing in Data Warehouses. Foundations of Management, 2(1), 99-116.

Saunders, M., Lewis, P. & Thornhill, A. (2009). Research methods for business students. 5th Edition, Pearson Education.

Sarawagi, S., Agrawal, R., & Megiddo, N. (1998). Discovery-driven exploration of OLAP data cubes. In International Conference on Extending Database Technology (pp. 168-182). Springer Berlin Heidelberg.

Sasisekharan, R., Seshadri, V., & Weiss, S. M. (1996). Data mining and forecasting in large-scale telecommunication networks. IEEE Expert, 11(1), 37-43.

Saxena, R., & Srinivasan, A. (2012). Business analytics: a practitioner's guide. Springer Science & Business Media.

Schläfke, M., Silvi, R., & Möller, K. (2012). A framework for business analytics in performance management. International Journal of Productivity and Performance Management, 62(1), 110-122.

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. European Journal of Information Systems, 23(4), 433-441.

Shen, L. Z. , Liu, S. H. , & Chen, S. K. (2012). Design and implementation of ETL in police intelligence data warehouse system. Applied Mechanics and Materials, 121 , 3865–3869.

Simoudis, E. (1996). Reality check for data mining. IEEE Expert: Intelligent systems and their applications, 11(5), 26-33.

Spekman, R. E. (1988). Strategic supplier selection: Understanding long-term buyer relationships. Business horizons, 31(4), 75-81.

Tellis, W. M. (1997). Application of a Case Study Methodology. The Qualitative Report, 3(3), 1-19.

Thalhammer, T., Schrefl, M., & Mohania, M. (2001). Active data warehouses: complementing OLAP with analysis rules. Data & Knowledge Engineering, 39(3), 241-269.

Theodorou, V., Abelló, A., Thiele, M., & Lehner, W. (2014). A framework for user-centered declarative etl. In Proceedings of the 17th International Workshop on Data Warehousing and OLAP (pp. 67-70). ACM.

Tyagi, S. (2003). Using data analytics for greater profits. Journal of business strategy, 24(3), 12-14.

Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. EDUCAUSE learning initiative, 1, 1-11.

Wadhwa, R. S. (2014). Framework for Product-Mix and Changeover Manufacturing Flexibility. International Journal of Computer Science Issues (IJCSI), 11(2), 223.

Wang, T. , Hu, J. , & Zhou, H. (2012). Design and implementation of an ETL approach in business intelligence project. Advances in Intelligent and Soft Computing, 124 , 281–286.

Watson, H. J., Goodhue, D. L. & Wixom, B. J., (2002) The benefits of data warehousing: why some companies realize exceptional payoffs? Information and Management 39(6), 491–502

Wiersema, M. F., & Bowen, H. P. (2008). Corporate diversification: The impact of foreign competition, industry globalization, and product diversification. Strategic Management Journal, 29(2), 115-132.

Winter, R., Bischoff, S., & Wortmann, F. (2011). Revolution or evolution? reflections on in-memory appliances from an enterprise information logistics perspective. In IMDM (pp. 23-34).

Wixom, B. H., Yen, B., & Relich, M. (2013). Maximizing Value from Business Analytics. MIS Quarterly Executive, 12(2).

Wixom, B. H., and Watson, H. J., (2001) An empirical investigation of the factors affecting data warehousing success. MIS Quarterly 25(1), 17–41.

Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence systems. Journal of computer information systems, 50(3), 23-32.

Zaghloul, M. M., Ali-Eldin, A., & Salem, M. (2013). Towards a Self-service Data Analytics Framework. International Journal of Computer Applications, 80(9).

Zimmer, M., Baars, H., & Kemper, H. G. (2012). The impact of agility requirements on business intelligence architectures. In System Science (HICSS), 2012 45th Hawaii International Conference on (pp. 4189-4198). IEEE.

Zode, M., (2007). The evolution of ETL, from hand-coded ETL to tool-based ETL. Cognizant Technology Solutions.

Zuboff, S. (1985). Automatefin-fonnate: The two faces of intelligent technology. Organizational Dynamics, 14(2), 5-18.