Aalto University
School of Engineering

Degree Programme in Mechanical Engineering

Erno Bister

The business intelligence transformation – A case study research

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science (Technology)

Supervisor: Professor Kari Tanskanen

Thesis advisor(s): D.Sc. Timo Seppälä

Aalto University
School of Engineering

Aalto University,
P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Erno Bister

Title of thesis The business intelligence transformation – A case study research

Degree programme Mechanical Engineering

Thesis supervisor Professor Kari Tanskanen

Code of professorship TU-22

Minor Industrial Economy

Thesis advisor D.Sc.Timo Seppälä

Date 3.3.2015

Number of pages 64

Language English

Abstract

The business intelligence (BI) field is changing rapidly. Companies need to be able to refine information from data faster than ever before to stay competitive. This is not an easy task in a world where volume and variety of data is increasing constantly. Social media and devices produce data on a continual basis. The BI market has responded to this challenge by providing self-service platforms that require less technical proficiency. These new tools enable business users to create and refine information more conveniently in cooperation with company IT.

As new technologies empower business users and relieve the workload of BI developers, new challenges arise. Managing information flows and data models in an environment with increased number of citizen developers, number of source systems and volume of data requires a clear understanding of related challenges.

This work examines a BI culture change in a Finnish energy distribution company divested from a larger energy company in 2014. Effects of moving from IT administered BI development closer to business driven development are examined through a case study research.

The case study research identified the cross organizational knowledge, overlapping roles and low technical proficiency requirements of the BI platform as the key components of the improved ability to create BI solutions for business needs. The findings are in line with critical success factors proposed by BI related publications.

Keywords Business intelligence, big data, self-service, case study research

Aalto-yliopisto
Insinööritieteiden korkeakoulu

Aalto-yliopisto, PL 11000,
00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Erno Bister

Työn nimi Liiketoiminnan analytiikan murros – Tapaustutkimus

Koulutusohjelma Konetekniikan koulutusohjelma

Valvoja Professori Kari Tanskanen

Professuurikoodi TU-22

Sivuaine Teollisuustalous

Työn ohjaaja D.Sc.Timo Seppälä,

Päivämäärä 3.3.2015

Sivumäärä 64

Kieli Englanti

Tiivistelmä

Liiketoiminnan analytiikassa (Business intelligence) eletään murrosvaihetta. Yritysten tulee kyetä jalostamaan tietoa nopeammin kuin koskaan aikaisemmin säilyttääkseen kilpailukykynsä. Sosiaalinen media sekä sensorit ja laitteet generoivat dataa valtavalla nopeudella sekä vaihtelevissa formaateissa. BI-markkinat ovat reagoineet muutokseen tarjoamalla erilaisia käyttäjäläheisiä tuotteita, jotka vaativat aikaisempaa vähemmän teknistä perehtyneisyyttä. Tuotteet mahdollistavat analytiikan tuottamisen lähempänä varsinaista liiketoimintaa yhteistyössä IT:n kanssa.

Kehityksen muuttuminen käyttäjälähtöisemmäksi tuo kuitenkin mukanaan uusia haasteita. Informaatiovirtojen ja datamallien hallinnointi ympäristössä, jossa on aikaisempaa enemmän käyttäjiä, lähdejärjestelmiä sekä dataa vaatii tarkan ymmärryksen aiheeseen liittyvistä haasteista.

Tässä työssä tutkitaan BI-kulttuurin muutosta suomalaisessa sähkönsiirtoon erikoistuneessa yrityksessä, joka myytiin omaksi yritykseksen 2014. Kehityksen muuttumista IT-lähtöisestä enemmän liiketoimintalähtöisemmäksi tutkitaan tapaustutkimuksen avulla.

Tapaustutkimus tunnisti organisaatiorajat ylittävän tiedon, yhteiset vastuut kehitystyössä sekä BI-alustan helppokäyttöisyyden tärkeimmiksi tekijöiksi, jotka paransivat yrityksen kykyä toteuttaa liiketoiminnan analytiikkaa. Löydökset olivat linjassa aikaisemman aihepiiriin liittyvän kirjallisuuden kanssa.

Avainsanat Business Intelligence, big data, self-service, tapaustutkimus

Preface

I became acquainted with business intelligence originally by a sheer coincidence. As a student of mechanical engineering I received a summer job offer from Caruna after three years of studying. My tasks consisted of monthly spreadsheet reporting of network building projects. At the time I had very little understanding over distribution business and business analytics. Over the years my managers have allowed me to strengthen my technical and analytical capabilities while performing daily tasks and routines. After three years the interest towards business intelligence exceeded my interest in mechanical engineering. At this point the writing of master's thesis became relevant. Caruna offered me a possibility to write my master's thesis to support the business intelligence strategy definition. I gladly accepted the challenge and I was able to write the thesis to my minor being industrial engineering and management.

I wish to thank all the managers who have allowed me to learn and develop my skills regarding business intelligence: Pauli Miettinen, Arto Liikanen, Mikael Mickelsson and Harri Hauta-Aho.

Special thanks goes to my thesis advisors Timo Seppälä from the Department of industrial engineering and management and Heikki Linnanen from Caruna who contributed their valuable time as advisors to this work.

I also wish to thank my father Timo Bister for providing guidance during the last steps of finishing this work.

In Espoo, January 2015

Erno Bister

Contents

List of figures

List of tables

1 Orientation to the field of research

Introduction

Business intelligence (BI) conventions are in change. Company analytics and reporting practices have been moving from IT administered and technology heavy solutions to more business oriented self-service platforms. Companies thrive to achieve a managed single version of truth of company performance without long lead times of traditional BI and data warehouse development projects. Analytical tools are provided to the hands of the business users to enable them to apply business rules to analytic applications without the need to communicate them to a technology oriented IT developer. While providing users with more freedom, the BI community is trying to avoid falling in to the spreadsheet hell with the new technologies, while still gaining extra agility that the self-service technologies provide.

Company analytical needs are not satisfied any more by providing additional tools to analyze company internal transaction processing system data. Companies seek a 360 degree view of the company performance combined with external market data that enables a more accurate base for forecasting and decision making. Increased number of data sources increase the amount of data. This requires new practices to manage the logical data models in addition to the raw computing power that the utilization of big data requires.

To achieve all this, the business intelligence systems need to be systematically planned and implemented within the company IT and process environments. The next generation of BI systems is designed for users with less experience in working with analytical tools and data models. The business users, on the other hand, need to develop skills to utilize the new tools.

In this work we establish a business intelligence deployment framework using literature review. In addition to the literature review, a case study research is presented. The case study evaluates two different business intelligence report development projects regarding the same source system. The projects were carried out using different tools in different technical and organizational environments. The work is conducted in a major Finnish utility company that was renewing its core IT systems after becoming an independent company as a result of divestment process in the early 2014. The framework and findings from the case study research will be utilized when creating the company business intelligence strategy.

Backround

The case study in this work examines the effects of change in business intelligence practices during divestment of Caruna. Caruna was part of a major Finnish power and electricity distribution company Fortum before the divestment in early 2014 when it was separated to its own company. Fortum sold all of its distribution businesses that were located in Finland, Sweden and Norway. The Finnish distribution business was divested to a standalone company that became Caruna. Caruna owns 20% of Finnish local distribution network and delivers electricity to 648 000 customers.

The case study evaluates the differences in development times, costs and quality of business intelligence systems before and after the divestment from the standpoint of Caruna. The BI development during Fortum era was heavily characterized by centrally managed IT department that offered services to all Nordic country organizations. The BI department was highly structured with strictly defined roles, responsibilities and processes that were in line with management approaches suggested in literatures such as Kimball & al. 2008 and Moss & Atre 2003. Fortum utilized data warehouses that gathered information from multiple source systems and provided data for analysis and reporting.

The divestment of Caruna forced the company to establish its own IT infrastructure including the BI systems. Due to the dissatisfaction towards the previous BI solutions, the BI tools and practices were among the first in the change list. Caruna started using lighter tools compared to massive data warehouse implementations that required less technical proficiency and enabled development of reports and applications closer to the business functions.

Statement of the problem

The purpose of this work is to evaluate and identify factors and practices that enable successful BI management in a long run. During the time of writing this work the majority of core business systems of Caruna were still legacy systems from Fortum. Projects to implement ERP systems, grid control systems and multiple smaller systems were well on their way. Caruna's reporting and analysis practices had been heavily characterized by manual spreadsheet reporting by multiple people within the organization. This resulted in different versions of same metrics and due to erroneous nature of manual work metrics showed some inconsistency between reports even if the author remained the same.

Decentralized manual spreadsheet reporting was largely a result of corporate IT's inability to fulfill information needs with more sophisticated automated tools. People had been used to very long lead times of BI projects with difficulties regarding communication with developers. Spreadsheet reporting as an alternative provided more control to the contents of the reports and shorter development times due to the less complex nature of the technical solution and less time spent on communication with developers. More agile online analytical tools that provided automated updates from databases were taken into use during first quarters after the divestment. Initial feedback regarding user satisfaction was significantly higher than before, and development times were greatly shorter. All this was achieved with less costs than before. The success of agile approach to automated online reporting, however, was mainly achieved in a small part of the organization, which required less connections to data sources and was manageable with few key people. The ongoing renewal of all business applications and systems was an effort becoming more and more complex in terms of data flows and models. To cope with all that Caruna had to establish clear principles of how to incorporate new data and information to the BI system.

Research design and methodology

I had been working for the company for four years during the time of writing of this work. My roles and responsibilities evolved from spreadsheet reporting tasks to the main developer of the company business intelligence system. Practically I am playing two roles regarding this work. Partly I am examining the matter as a thesis writer and partly sharing what I have been able to identify from the work of my own and from colleagues and experts during the BI culture change.

The work is divided into two sections consisting of literature review and case study that examines perceived phenomena against the practices proposed by literature. The literature review leans heavily to traditional data warehousing methods, whereas the actual practices in use in the company are highly interconnected with the new technology that the market provides.

Validity, reliability and limitations

The case study is conducted within a company that possesses a natural monopoly. The market is also very predictable and stable since consumers and businesses are going to use electricity and therefore require distribution services in nearly all imaginable scenarios. This limits information needs and business intelligence solution requirements from a market intelligence perspective. The case study, however, concentrates more on BI solution implementation and its ability to meet business needs. In that sense, the results should be applicable in wide range of different companies.

The dualist role of the author as an employee and as a researcher may affect the observations to some extent. This should be kept in mind when assessing the credibility of BI development success evaluation where the author was more responsible of the solution developed to Caruna than in the solution developed by Fortum corporate IT. The author's experience in BI field extends over 1.5 years of time within the single organization, which affects the completeness of the understanding of BI field.

Structure of this work

This work is structured to five main chapters. The first chapter is an introductory chapter that briefly describes the modern business intelligence field, provides research question with description of used methodologies and provides background of the electricity distribution company called Caruna where the case study research takes place. The second chapter introduces terms related to information, how businesses are refining and utilizing information for their competitive advantage, and what is the role and current state of business intelligence in this. The third chapter is a literature review that walks the reader through a business intelligence deployment from beginning to implementation and serves as a background for the case study research. The fourth chapter presents a case study research that evaluates two report implementation projects regarding the same source system within an electricity distribution company before and after it was divested from a larger Finnish energy and distribution company. The fifth chapter draws conclusions from the case study research and literature review and compares the findings from the two.

2 Information as an asset

In literature information is often accompanied with terms data and knowledge. Although these three are often used as synonyms in conversations they hold a slightly different meaning. (Jashapara 2004, Bocij & al. 2008) In order to discuss the topics such as information management and business intelligence, we need to establish these terms more precisely.

Data

Data is information without context or purpose. For example '01.01.2016', Helsinki, 54 and €304.53 are examples of data. By themselves they do not provide any meaningful insight to a person viewing them, and thus, they have little or no value. Unrelated pieces of data with no connection or significance are often described as noise. There are several definitions of data in use:

1. A series of non-random symbols, numbers, values or words,

2. A series of facts obtained by observation or research and recorded,

3. A collection of non-random facts,

4. A record of an event or fact.

To become meaningful data needs to be transformed, processed and presented accordingly. When the steps mentioned above are in place, data becomes information.

Information

Jashapara 2004 uses the definition 'systematically organized data' about information. Information that is provided at the right place, at the right time and in an appropriate format enables organizations to make decisions and work towards a shared goal. To transform data into information it needs to be placed within a context. The data examples given in a previous paragraph do not enable any insight on their own. However, if a person needs to travel to Helsinki on 01.01.2016, a record in an airline company's web page telling that there is a flight number 54 flying to Helsinki on 01.01.2016 leaving from terminal 2 at a price of €304.53, these data together, within a context provide information to this person that he can use to organize his trip. Jashapara also states that in order for data to become information it has to have meaning, relevance and purpose. Meaning in this context cannot be evaluated in scientific or absolute terms but rather as a meaning that the receiver gives to the data. If the recipient does not know the context the data is taken from, it is still data. Therefore the distinction between data and information is subjective.

Data can be systematically processed to information. Chaffey & White 2011 present the following stages of the transformation process:

1. Capturing the data,

2. Routing the data to location for processing,

3. Processing the data to produce information,

4. Distributing information to users,

5. Analyzing and interpreting the data by users coupled with their skills and previous experience to take actions.

Activities that take place during the transformation process are placing data to context, performing calculations with data, classifying data and aggregating data to higher level e.g. total sales by country. (Chaffey & White 2011)

Knowledge

Knowledge can be described as tacit or explicit. Tacit knowledge refers to person's ability to know how to carry out certain tasks. Explicit knowledge is closer to a traditional view of knowledge where certain information provides understanding of a certain matter. (Bocij & al. 2008)

Tacit knowledge is hard to capture, transmit or describe accurately. It includes skills like knowing how to act and respond in social situations, knowing which goals benefit the organization more than others and understanding how seemingly disconnected pieces of information relate to each other. The experiences, values, possessed information, skills and background determine which kind of tacit knowledge a person has. Explicit knowledge can be documented and captured more easily. It is more structured, it can be transferred using documents or taught. Examples of explicit knowledge are instruction manuals and meeting memos. (Jashapara 2008, Bocij & al. 2008)

The multidisciplinary nature of knowledge management results in varying definitions of the subject depending of the presenters' viewpoint. Some definitions have their roots more in information sciences and others lean more to human resource perspective. (Jashapara 2008) Bocij & al. 2008, however, state that many researchers agree that it involves utilizing a person's personal abilities, such as his or her experiences, intuition and analytical skills.

The goal of knowledge management is to increase intellectual capital and enhance organizational performance. This view point emphasizes the strategic perspective of knowledge management. (Jashapara 2008)

During previous paragraphs we established basic understanding over information fundamentals. During next paragraphs we are going to explore how information use plays a role in running businesses.

Business information systems

Most modern businesses are heavily driven by information. Technologies have evolved significantly and supplying business with required information is cheaper and more flexible than ever before. Information is accessed with mobile devices in addition to traditional computers and systems are required to work seamlessly together. The customers, workers, manufacturing sites and transportation routes are dispersed all over the world and a company is required to manage its business on an ever increasing speed to remain competitive. Internal information needs account for stock levels, distribution of costs and sales along with visibility to the supply chain and to ongoing projects. External information enables companies need to stay up to date with market preferences, competitors' prices and the supply and cost of finance. The amount of information, to address all the business needs today, is so vast and constantly changing that it is hard to imagine that it is even possible to run a multi-location business without efficient use of technology. Technology is used to gather, store and view the information to support running the business and determining how it could be improved in the future. (Graham & al. 2008, Gendron 2014)

In order for a company to be successful the information systems (IS) need to be developed and integrated to the daily operations in a manner that supports fulfilling the company strategy. Information systems by their own do not help selling any products or provide analytical insights for doing so if they are apart from the actual business. Ng & al. 2013 present a model that shows the high level activities of a business, and how IS are a vital part of fulfilling and reviewing company strategy on a continual basis. The model is presented in figure 1.

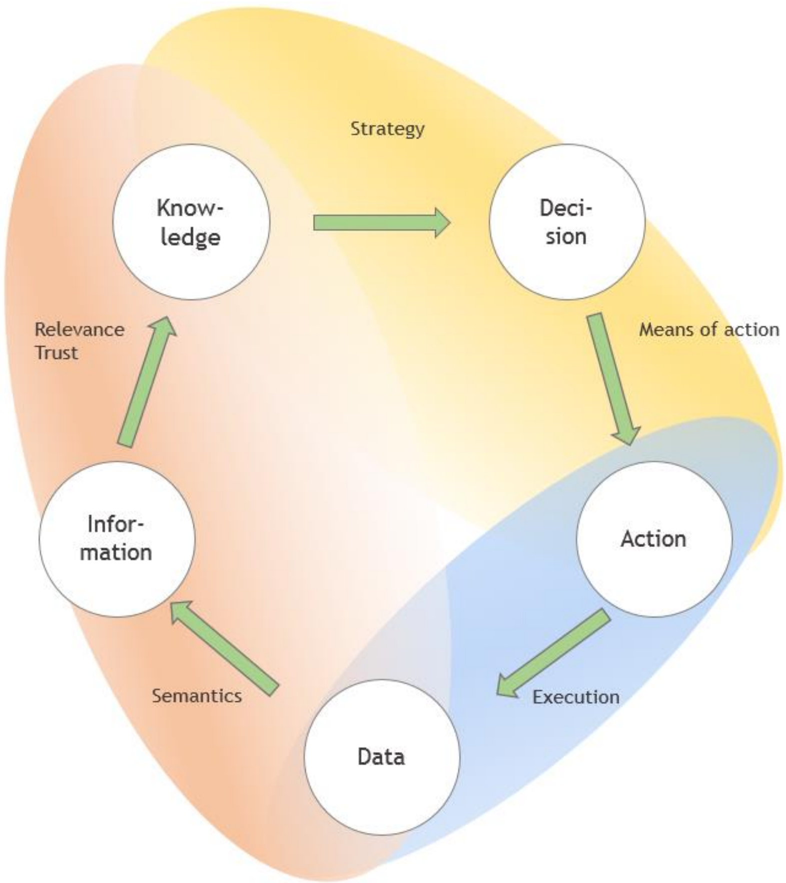

Figure 1 Business information flow (Ng. & al. 2013)

The key areas of the model are people, transactions and BI/analytics. The role of people in business is to develop a strategy, make decisions that further the company strategy and act according to the decisions made. Taken actions are recorded as transactions to the company information systems. An example of this could be that a batch of goods was ordered from supplier A to a company in location B on a certain date. Recording actions on a daily basis generates data that can be refined into information. By refining the data it is possible to gain information whether or not the organization has been able to act accordingly towards the goals defined in the strategy. After this cycle the company executives have the opportunity to review the strategy and decide if to continue on the same track or direct efforts somewhere else. Depending on the novelty of information systems and related business processes the quality of the information at the end of the cycle may vary considerably. At its best the executives have the information of the exact problems that hindered the result and the reasons behind successes. At its worst the company might realize that some data was generated but no meaningful information was possible to refine from it. (Ng. & al. 2013) In these two extremes the tools that the executives have to guide their companies to success are far apart from each other. A perfect integration of IS to business in reality is, however, nearly impossible, and often companies have to work with solutions that are good enough, cost effective and provide the most important functionalities.

After having established the basic information production and use cycle we are going establish the basic understanding over business intelligence, explore the lifecycle of business intelligence as a phenomenon and present some current trends.

Business intelligence

The term business intelligence emerged when an IBM researcher Hans Peter Luhn described BI as "the ability to apprehend interrelationships of presented facts in such a way as to guide action towards a desired goal." (Luhn 1958) The definition still holds true although the nature of available data and tools have changed profoundly. Business intelligence is gaining more and more attention and is no longer seen as a side activity that only supports businesses. Whole new businesses emerge that utilize available data and try to seek competitive advantage over competitors by moving from capital and asset heavy models to information driven business models. (Ng. & al. 2013, Schmarzo 2013, KPMG 2015) A KPMG report from 2015 conducted a questionnaire with 260 participants of which 130 were senior institutional investors and 130 senior decision makers at sell-side firms. The findings of the questionnaire present that 37% of the participants believe data and analytics is currently changing competitive dynamics, and 57% of the participants believe it will do so within the next three years. The questionnaire also reveals that investors are examining company BI strategies carefully before purchasing decisions. Companies with little or no attention to analytics were seen as less interesting and otherwise promising investments were ruled out due to a lack of proper BI strategy. Reshi and Khan 2014 similarly state that BI implementation is the most important implementation that can help to gain competitive advantage.

Business intelligence seeks to offer timely and accurate information for decision makers to enable them to make better decisions. It is used to identify changes in the market and determine, how the company could position themselves to utilize the existing market. Internally BI is used to analyze company performance and identify actions that would enhance company productiveness and reduce costs. (Khan & Quadri 2014, Kimball & al. 2008)

BI analysis spectrum introduced by Loshin 2013 can be used to understand different analytical activities to which BI systems try to offer tools. The less advanced analysis techniques are on the left and as the spectrum extends more to the right, the techniques become more sophisticated and require more maturity from the organization and BI systems. The analysis spectrum is presented in figure 2.

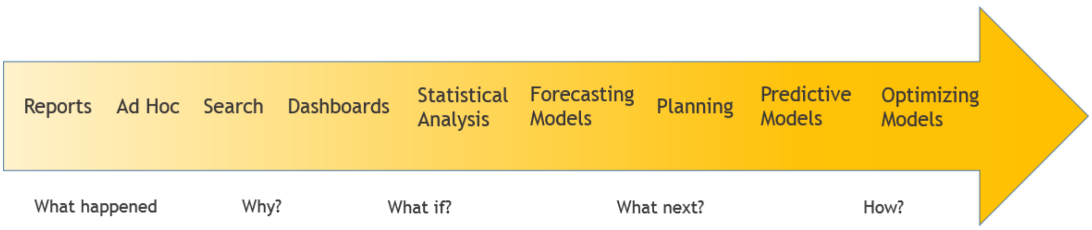

Figure 2 The BI analysis spectrum (Loshin 2013)

Traditional business analytics have focused on delivering reports periodically to assess the performance of a company in the past. The reports have tried to answer questions like "How did we perform during the first quarter of the year?" The reports were traditionally generated by running queries against multidimensional data warehouses, where the data had been extracted for the purposes of analysis. The next maturity level of an organization using BI system is where people have the tools and willingness to explore causes behind the past performance. Basic BI concept that supports further exploration of data from a summarized level is called drilling down. Basically this means that a user has the immediate access to row level transactions and accompanied attributes that sum up to the company or organizational level metrics. As analytical capabilities become increasingly real-time, the next step is to start looking into the future with analytics instead of just analyzing the past. Forecasts are a mix of historical data, statistical analysis, simulations and predictions that try to picture future business landscapes and how a company should function in them. (Ng. & al 2013, Khan & Quadri 2014, Loshi 2013)

The increased interest and demand of improved business analytics evokes new BI vendors and technologies that try to answer the needs of business users and analysts. Self-service platforms have been developed to empower business users to create their own analysis's to shorten the development time of BI applications. (Gendron 2014, Ng. & al. 2013, Krawatzeck 2015) Big data centric companies are in the forefront of developing distributed file systems and analytical platforms that are able to handle massive data amounts in real time. Since big data and self-service platforms are highly dominant in publications and BI product releases during the time of writing this work, they are both introduced in their own sub chapter.

Big data

The amount of data in the world has exploded within last few years. IT technologies become more and more accessible to consumers and enterprises each day as technologies evolve and prices decrease. The increased number of people using computers, mobile devices and internet services leads to ever increasing speed that data is generated. Being able to utilize the massive amounts of data may lead to competitive advantages and create new business opportunities that have not existed before. People using services like Facebook, LinkedIn, Twitter and Google perform vast amount of selections, likes and clicks every day. By storing and analyzing this consumer behavior new business opportunities emerge. People can be targeted with tailored advertisements while they surf their way on the internet, and different bundles can be offered while they are checking out with their order in an online shop. In addition to this, geospatial data generated by mobile devices that people have always with them add a piece to the picture of a consumer profile. People can be targeted with advertisements while they pass a shop that might be of interest to them according to previous consumer behavior. Sensors and machines generate real time data at production facilities and remote locations about their current metrics. Preventive maintenance programs can be created by identifying patters from data that precede a breakdown. By utilizing that information it is possible to predict breakdowns in the future and repair them before it is too late. (Schmarzo 2013 and Morton & al. 2014)

This phenomenon is often referred as big data. There is no single definition for big data but it is often characterized with some of the following according to Schmarzo 2013, Morton & al. 2014 and Ng & al. 2013:

1. **Volume**, the amount of data requires special considerations to be stored and analyzed,

2. **Variety**, the data consists of multiple types of data of which some is more and some less structured. Data potentially comes from multiple sources and may include photographs, different types of documents or other form of binary data,

3. **Velocity**, the data is produced in such speed that it utilizing it requires special attention. Data ages quickly and operating with older data is not valuable.

Out of these three the variety is a characteristic that stands most apart from traditional data storing. Whereas relational databases are strictly defined with datatypes and connections with certain elements, big data may contain highly unstructured data that is not necessarily designed to fit neatly with other data. This data may include videos, pictures, text documents, conversations and such. Relational databases do not scale well with increased volumes of data and they have limited capabilities to deal with different data types. To answer this new demand of storing and analyzing unstructured data new NoSQL databases emerged. NoSQL can mean Not-only-SQL or No-Never-SQL depending on the database product. The principal categories of NoSQL databases are key-value stores, document stores, extensible record stores and graph databases. All these databases try to offer a better way to manage some aspect of big data e.g. graphical databases are best for analyzing different connections of data that might be difficult to analyze otherwise. (Morton & al. 2014)

In addition to different databases, some new distributed computing technologies had to be developed to be able to benefit more from the data explosion. Data intensive companies like Amazon, Google, Yahoo! and Facebook were first in trying to create better tools to utilize massive amounts of data. Their value proposition is built around collecting huge data volumes and later monetizing that data. Therefore they had the biggest incentive to find ways to manage all the data. Their effort ultimately led to technologies we now know as MapReduce and Hadoop. (Schmarzo 2013, Morton & al. 2014)

Apache Hadoop

For many Apache Hadoop has become synonymous with big data. It is an open-source software framework that supports natively parallel and distributed applications. It is designed to run multiple applications on large clusters of commodity hardware. A key component of Hadoop is the MapReduce paradigm. In MapReduce the application is divided in smaller fragments of work which can be executed on any node in computer cluster. Hadoop also provides a Hadoop distributed file system (HDFS) that stores data to different computer nodes. Both MapReduce and HDFS are designed in a way that automatically handles node failures. It enables applications to work with thousands of computation-independent nodes and petabytes of data. The entire Apache platform is considered to consist of Hadoop kernel, MapReduce, HDFS and related projects including Apache Hive, Apache HBase, and others. (Schmarzo 2013, Morton & al. 2014, Loshin 2011)

Self-service BI

A traditional business intelligence approach has often been heavily driven by IT department that delivers information to business users through dashboards and periodically generated reports. Being IT driven, BI projects have often had difficulties in answering the information needs of the users. One reason for this is the low level of business understanding among the BI developers. (Kimball & al. 2008) The technology nature of BI development has been heavily relying on massive data warehouse implementations with complex logics and very thorough documentation. Long development times in a world of constantly changing businesses and needs accompanied with information gap between developers and users have together resulted in a low success rate on BI implementations. Even if the BI development project would mainly meet their goals within the organization, the users are likely to have more information needs and ideas for new projects than the IT department is ever capable of delivering. This is often referred as applications backlog. (Bocij & al. 2008)

As an answer to information gaps and long development times, the BI industry has started to offer advanced self-service tools for business users. Business users are able to create their own reports which are connected to the data sources and provide updated real time information. The industry trend is to provide analytical tools to a larger number of people within the organization and break the image that BI tools are only meant for analysts and IT experts. (Ng. & al 2013) Spreadsheet reports are a traditional and a very easy to learn tool for users to create their own reports. Many transaction systems are able to provide a data export in a form of single table. Users can use the data and create their own analyzes and models and share them with their colleagues through e-mail. A problem related to spreadsheet, however, is that it allows almost limitless freedom to modify and calculate the data. This may often result in multiple versions of same calculations that do not match in result. Another pitfall of using spreadsheet is the lack of version control. (Bocij & al. 2008)

Modern self-service BI system is a mix of freedom and centrally managed platform that ensures up to date relevant and correct data for the users. It enables users to add their own data possibly from external sources (e.g. currency information, weather information) and analyze possible connections between internal and external data. BI systems need to be designed for people that have less experience and training on analytical tools. Therefore data needs to be easily available, the user interface needs to be self-intuitive and the application itself cannot be overly complex. (Ng. & al 2013)

Chapter summary

In this chapter we have discussed basic terms related to information and established an understanding how data is captured and refined into information within a company. After that we delved more into the subject of business intelligence and discussed where the term originates from and how the practice has evolved to its current form. We also discussed topics of big data and self-service BI due to their weight in BI related conversation and in market during the writing of this work. The next chapter walks the reader through business intelligence system implementation and development. The overall view presented in this chapter serves as a high level map which can be used to place more detailed activities and concepts to a broader context.

3 Business intelligence deployment framework

In this chapter we will walk through a business intelligence system implementation from early assessments to post-deployment governance. Martinsons 1993 states that "No single model exists for creating a strategic business intelligence system. The mission and activities of the organization, and the style of senior managers, define its specific objectives and structure." In addition to differences in desired goals, starting points and capabilities to carry out such an endeavor vary depending on the maturity of the information systems and organization of the company. A business intelligence implementation is not a project according to a traditional specification (Olszak & Ziemba 2012). An operational system implementation can be considered ready when the system is handed over to the organization using it. Bug fixes and reconfigurations might follow the handover but the project is intended to meet its completion criteria as well as possible and then continue into production. BI implementation serves as a starting point for continuous development where the BI system matures and the organization develops new skills and ways of utilizing the new technologies. (Yeoh & Koronios 2010) A BI analysis spectrum by Loshin 2013 was presented earlier in chapter 2. At the left end were reports and adhoc analyses and at the right end predictive and optimizing models. Being able to identify what is happening currently and in the past, is a good goal for an initial BI implementation. The ultimate goal, however, is to set foundation in way that supports incremental development cycles that gradually raise the capabilities of the BI system to support forecasting and planning.

BI implementation involves numerous steps that include analysis of current systems and business needs, different design and development activities and hardware and software acquisitions. The use of BI within a company should aim at continuous evolvement of the system by evaluating past projects and fueling next steps with knowledge gained from successes and failures. Gangadharan & Swami (2004) divide high level activities of a BI system lifecycle in analysis, designing, development, deployment and evolution. They present the activities in a cyclical form where the preceding step leads to next step and when one cycle is completed the next one begins. The lifecycle of BI system is depicted in figure 3.

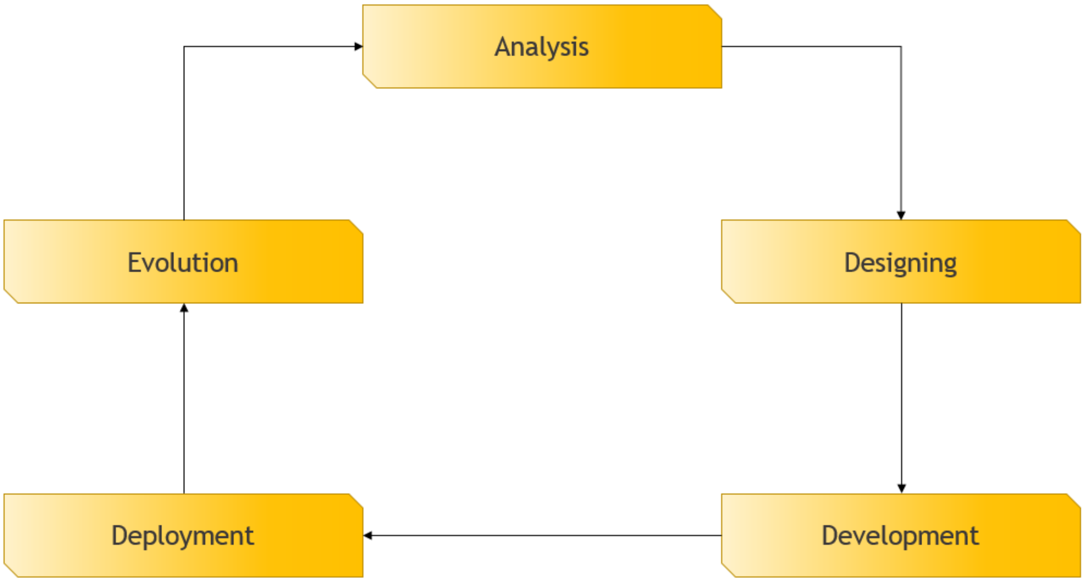

Figure 3 Business Intelligence system lifecycle (Gandrahan & Swami 2004)

The deployment framework in this work also presents required activities in a sequential order. Due to the nature of BI projects the activities are, however, highly overlapping and conducted multiple times during the project with different level of detail. For example, the business case evaluation and gathering business requirements serve as a starting point and initiator of a BI system implementation. It is, however, unlikely that all the required information can be extracted with a single iteration of requirements engineering. The initial requirements engineering activity seeks to answer very high level design questions so that the BI architecture supports later data modelling and application development. The requirements gathering needs to be re-conducted during data modelling and application development phases to collect all the required details, user preferences, required update frequencies and other significant aspects that need to be defined in the solution. It is also difficult to answer precisely what is the return of investment (ROI) for a BI system. It is very hard to estimate how significant the business opportunity actually is before exploring the data with BI tools. In addition to the difficulty of evaluating the business opportunity, it is also difficult to evaluate how well the system is capable of leveraging and delivering the business opportunity in the end. Although the BI implementation includes a lot of uncertainty and all questions cannot be answered in the beginning, it is vital to understand related activities and try to justify the project in a best manner possible before starting to refine the vision of the deliverable and begin the required activities that lead to the deployment. (Kimball & al. 2008, Moss & Atre 2003)

BI implementation success factors

Kimball & al. (2008) suggest that before launching a BI project in a company the readiness of the organization for such a project or system should be carefully assessed. In literature the most influential factors that contribute to the success of a BI initiative are often called as critical success factors (CFS). (Yeoh & Koronios 2010 and Olszak & Ziemba 2012) Identifying the most significant factors for BI successes enable better assessment of how far a BI initiative should try to reach, which areas are in good shape regarding the project, and which require extra attention.

Yeoh and Koronios (2010) group the CFSs under three categories being organization, process and technology. Organization related factors include elements such as committed managerial support and strong business case. Process factors include team and competency issues and project and change management factors. On technology side data and infrastructure related questions are assessed. The CFSs by Yeoh & Koronios are presented in figure 4.

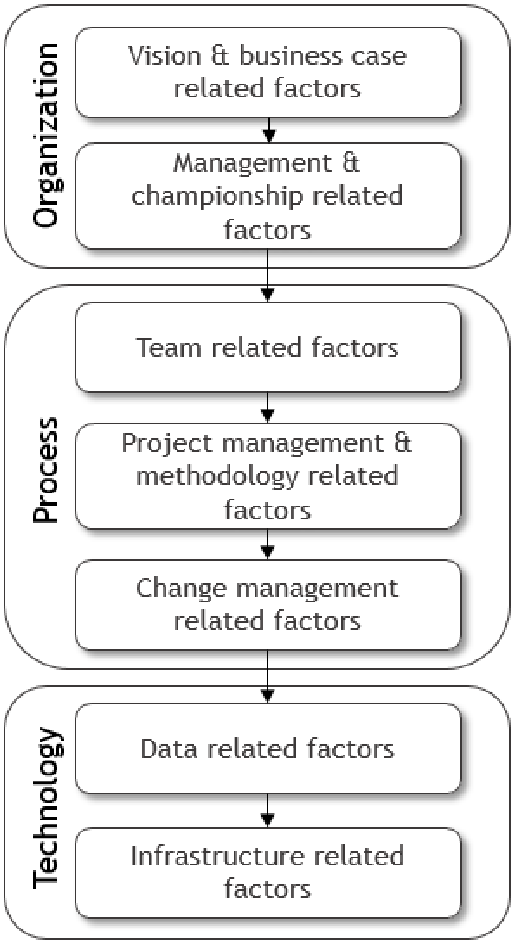

Figure 4 Critical Success Factors (Yeoh & Koronios 2010)

Out of all CFSs the committed management support is recognized as the most determining. This is emphasized by Yeoh & Koronios (2010), Olszak & Ziemba (2012) and Atre (2003). The senior business management have the best vision for the potential impact of a BI solution. Strong sponsors typically are willing to commit to the project if they see that it provides real value for the business. The support from senior management can help to overcome obstacles that arise from other CFS areas that could otherwise endanger the success of a BI initiative. Strong sponsorship makes it easier to secure required resources such as funding, human resources and other necessities. (Yeoh & Koronios 2010)

Another significant factor is that the project should be heavily business driven instead of IT driven. Systems built from IT perspective with little true knowledge about the actual business needs may often reach poor utilization rate. Atre (2003) lists ten major reasons that may lead to BI failure of which most of them include the viewpoint of being not enough business driven. Companies that successfully implement BI systems typically have an urgent need for improved access to information. Compelling business motivation may arise from a change in external business environment, internal crisis or the information needs created by recent acquisitions. In each case there is a demand coming from the decision makers and business community that can be satisfied with BI solutions. In simple terms, a need that can be satisfied with BI solution means a business case for BI solution. Real business cases provide real return on investment, and it is more likely that BI project is managed, supervised and carried out with more dedication if the benefits are clear. (Yeoh & Koronios and Alaskar & Efthimios 2015)

Kimball & al. (2008) do not raise the technical aspects to the three most determining factors. Their primary feasibility concern regarding BI readiness relates to the data itself. If the required data for needed information is not collected, faulty or overly complex, this creates a poor foundation for the actual solution building process (Kimball & al. 2008). Alaskar & Efthimios (2015) also conclude from various literature sources that data quality is one of the most determining factors of a successful BI system. In order to successfully create a cross-organizational BI system that connects data from various systems, the information must be extremely consistent and well integrated.

Gathering business requirements

In the previous paragraph we concluded that in order to be successful a BI implementation must be heavily driven by business. Simplistically it could be said that BI solutions ability to respond to business requirements is the single most important evaluation criterion when assessing the outcome of a BI implementation. If responding to business requirements is the main goal of a project, the correct identification of the business requirements heavily determines whether or not the solution will be used and accepted by the users. The business requirements determine the scope of the project, which data to model, which tools should be used, what analyses should be built and how the users need to be supported during and after the deployment. (Przybylek 2014, Yeoh 2011)

The gathering of business requirements occurs at two levels. On a higher level the business needs and priorities that influence the whole BI program are identified. This may cover the whole enterprise or the selected functions that the solution is aimed for. An example of a higher level business requirement would be the capability to support and integrate multi-location processes with multiple language selections. At the lower level the requirements may concern how information should be presented in certain reports or which metrics should be included in each report. (Kimball & al. 2008, Moss & Atre 2003)

Przybulek 2014 states that users rarely know exactly what they want from the BI solution. He proposes that the system analyst should make the users realize their needs. Kimball & al 2008 propose that instead of asking "What would you like to have in your report?" the developers should ask questions about every day work. What objectives do they have? Are there challenges that limit their ability to meet their objectives? Do they have visibility to the effects of their work and to the process? If the developers settle for requirements that the users are able to provide without any encouragement and facilitation, they might end up building the existing reports to a new platform. Martinsons 1994 points that often requirements engineering focuses on "what" questions instead of asking "why". Generally it is easier to gather information about current state of a matter but determining what has led to current state is more difficult. Finding reasons behind current state requires analysis and weighting of multiple possibilities. If a company is able to identify reasons leading to current state they have the ability to evolve and get better. (Kimball & al 2008)

Gaining a holistic picture of an organization’s information needs requires interviewing multiple stakeholders from different positions within the organization. Business executives are the company visionaries. They possess the information where the company is going and what future information needs are about to emerge. They also are aware of business pains that currently cannot be tackled with existing information resources. Executive level interviews mostly focus around strategic information and BI program level information needs. IT managers support the operational systems of different business areas. They know the capabilities of operational systems and are able to provide a current state view on to the company analytics solutions. They possess valuable information of BI application backlog and unfulfilled BI efforts. They are also able to tell how an improved BI environment could possibly relieve workload of IT by automatizing adhoc requests. IT staff work in close collaboration with business users and know the technical skills of business users and help to navigate among business people to find out who knows and what. Line-of-business managers deal with tactical decisions on a daily basis and are able to provide requirements to an application level based on their daily needs. Subject matter experts are the senior business analysts. They know the characteristics of the data better than anybody and are able to help with logical data modelling, calculation rules and data quality issues. Although the requirements engineering activities should focus on the needs of the business users, the IT personnel have the information about systems and what they are currently capable of and therefore provide possibility for a reality check with the collected business requirements. The requirements and capabilities should be matched in a balanced way that provides meaningful content in a sustainable and accurate manner. (Kimball & al. 2008, Moss & Atre 2003)

Kimball & al 2008 and Martinsons 1994 both suggest that face-to-face meetings and interviews should be the primary method in requirements engineering. Face-to-face meetings are interactive by nature and might reveal information that other participant might not realize to ask or tell without the interaction. Discussing requirements orally versus a requirements definition document through e-mail helps to confirm that the user and developer understand terms, needs and challenges the same way.

Project planning

After coming into a conclusion that a business case exists and the company has or is able to acquire the resources and culture to fuel a BI initiative, it is time to begin creating a more detailed project plan. A truly enterprise wide BI strategy requires years to mature. Development efforts and staffing must aim at long term improvement of company analytical capabilities. Factors like short term bonus incentives, willingness to jump into newest technologies, vendors looking to extend their contracts and consulting firms trying to sell more and more of their services create challenges for the long term project planning and realization. Project plan is a tool for aligning the efforts of the organization to serve in fulfilling the goal agreed. (Loshin 2013) One of the most critical project planning activities is defining the project scope. It is likely that during the requirements gathering far more needs emerge than it is reasonable to include into the initial project plan. Therefore it is necessary to define to which needs the project primarily tries to meet and which issues are left for later consideration. The project plan should include a schedule. What results can be expected at which timeframe and which activities are involved in the project? Locking down a delivery date and adjusting scope according to schedule is a safer way to go than keeping the scope fixed no matter what difficulties the project faces. The initial project scope might include all steps from current state to a tested and functional BI system that addresses most company's analytical needs. Another possibility for the scope is to establish physical architecture and incorporate most business critical system data to the BI system. Kimball & al. 2008 suggest that focusing on a single business process in the early phase of BI system development often yields good results since it narrows the complexity of the system. Delivering actual functionality to a single business area at a reasonable time also encourages management support more than a project that has been running for ages, and it will possibly deliver something in the future. The scope should be set jointly by IT function and business representatives to address both business value and technical aspects. (Kimball & al. 2008, Moss & Atre 2003)

The scope and schedule set the requirements for project resourcing. A BI solution consisting of few company internal operational systems requires minimal effort compared to enterprise wide implementation with multiple interfaces with vendor systems and capabilities of leveraging data from social media and other external data sources. More complex systems require more development time, personnel and hardware. Increasing the number of people working on the project may speed up the development to some extent but the returns diminish rapidly. Increasing the team size beyond the optimal may actually slow the process due to increased requirement of project coordination and communication. When scope, schedule and required resources are in place, it is possible to estimate the project costs and receive funding for the project. (Kimball & al. 2008, Moss & Atre 2003)

Technical architecture

When the analytical needs are identified through business requirements gathering and the high level project plan is in place, it is time to answer the question "How are we going to achieve the desired result?" Technical architecture is an overall plan about the BI system. It describes the data sources, data flows and transformation steps, data stores, application development tools and portals that share the information to users. The core functions of a BI system are to extract data from its original location, process it to a meaningful form and present it to the business users. A technical architecture plan serves as a blueprint for the system as well as a communication tool for team members and stakeholders to help them understand the complexity of the system and their own role in the process. An example of a BI architecture model is presented in figure 5. (Kimball & al. 2008, Moss & Atre 2003)

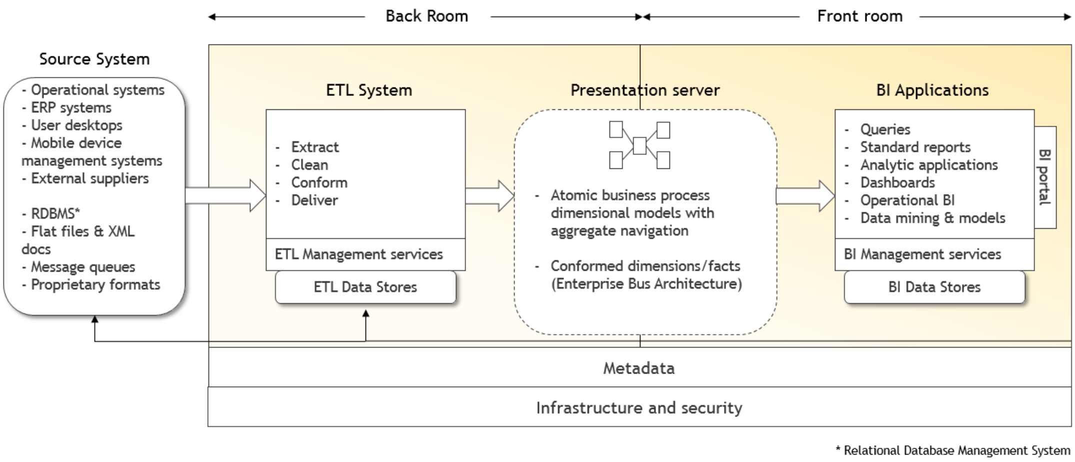

Figure 5 High level DW/BI system architecture model (Kimball & al. 2008)

The BI architecture can be divided into two sections - back room and front room. Back room represents processes that are hidden from users and serve as a foundation for front room operations. Back room is usually managed completely by IT since it requires an understanding of the technologies, procedures and relations of different activities. User rights must be managed to prevent unintentional changes that influence the operation of the whole system. Back room architecture involves connections to source systems whether they are in house operational systems, internal documents or external data sources. The goal is to load the data from its original location and bring it to a common platform through extract, transform and load (ETL) process. Extracting in this case means copying the data as it is in the original location to a staging area. Depending on the quality and readiness of the data for BI system the data goes through transformations that enhance the cleanliness of data (removing duplicate values, correcting data errors) and transform it to a form that serves the data model meant for reporting. The transformed data is then loaded to presentation server where it is available for the BI applications and analysis. The ETL process may vary greatly depending on the complexity of the data sets and data cleaning policies. Regardless of how thoroughly the ETL process is defined and designed the activity of loading data from source systems to a BI platform always occurs at some level in a BI system. The ETL process is discussed more in the data analysis and modelling section later in this work. (March & Hevner 2005, Dayal & al. 2009, Moss & Atre 2003)

The front room is more visible to the users with reporting applications and data made available for adhoc analysis with different BI tools. The users might have access to presentation server data model designs to be able to create their own queries and analyses depending on the solutions that are created for the business users. Applications serve pre-defined reports and queries that the business people can interact with and filter to meet their information needs. The front room activities are much more related to fulfilling the actual business needs. Although the backroom is not visible to the business users, no front room capabilities can be delivered, if the underlying processes and infrastructure are not in place and properly managed. (Moss & Atre 2003, Dayal & al. 2009)

Cloud vs. in-house

A company seeking to establish a BI architecture in 21st century should be aware of possibilities that utilizing cloud services can provide. Whereas maintaining servers and data in-house provides the highest level of control and security, cloud services provide flexibility, cost savings and high levels of on demand power. The increased flexibility is achieved with the cloud computing delivery model that can be described as "pay as you go". A client is no longer required to purchase dedicated servers for dedicated purposes, but can use the resources of a cloud provider on demand. This enables cost savings since the provider can offer massive computing power when needed but the client has to pay for hardware only when he is using it. (Gendron 2014, Gurjar & Rathore 2013)

Cloud resources are delivered to customers utilizing virtualization. By using virtualization, same hardware resource e.g. a server can be divided into virtual servers that seem to customers as they were actual servers. Customers are then able to work using the virtual server as if they had an own hardware server. Since multiple customers share the same resource the single customer naturally experiences less than full performance of a server unit. It is a rare case that all customers require the performance at the same time, and therefore the perceived performance compared to the cost of the server is significantly lower than of owning and maintaining an own hardware. (Gendron 2014)

Virtualization extends from simple resource delivery model. Cloud providers offer solutions where the customer has to be less and less aware of the details about how and where a certain service is maintained. This is a result of the emergence of software as a service (SaaS) models. SaaS separates the possession and ownership of software from its use (Turner & Brereton 2003). Software services are made visible for the users over the network enabling them to interact with the software product as if they were using it on their own machine. In SaaS the customer is not required to implement, host or maintain the software services they require. The application is hosted on the service provider's platform or a third party platform. The customer is able to log in to the application and start using it through the internet. (Gendron 2014) The term Business intelligence as a service (BIaaS) has emerged to describe BI related SaaS concepts (Chang 2014, Zorrilla & García-Saiz 2013).

The company must select the most suitable combination of cloud services and in-house solutions to achieve the desired level of flexibility, control and security. It is possible to store data in cloud and retrieve it to in-house server that hosts application development and publishing. Another option is to store data locally and use a BI visualization tool offered by a cloud provider to create reports and visualizations without a need of installing BI tools in-house. The cloud/in-house options are presented in figure 6. (Gendron 2014)

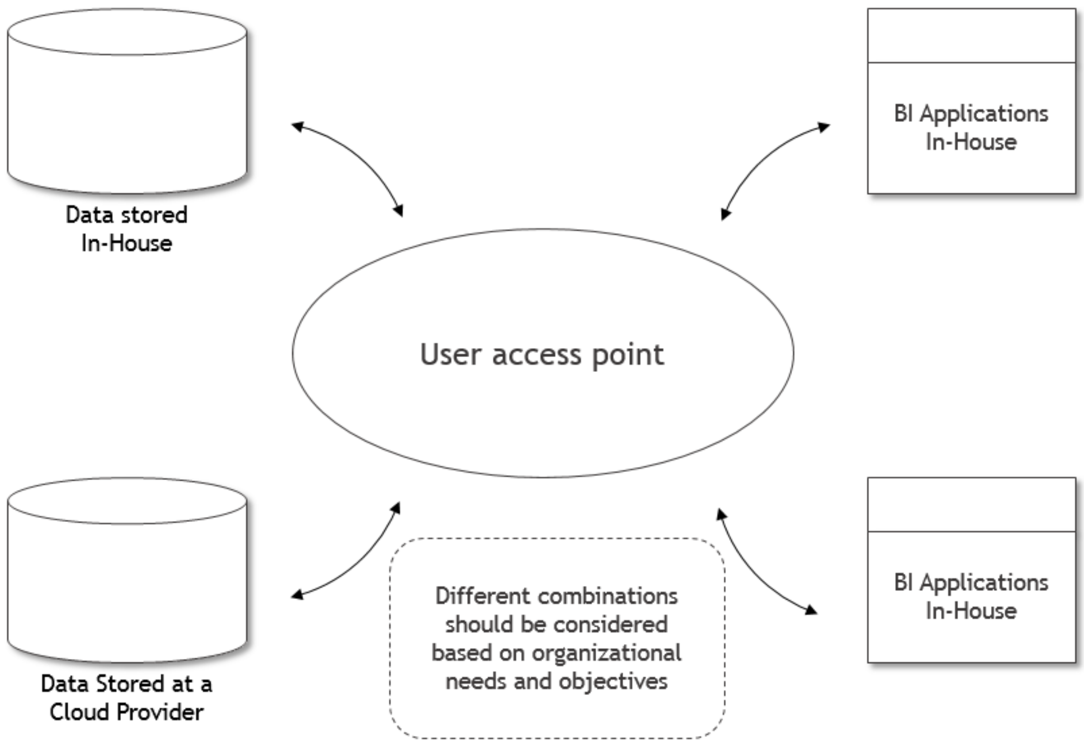

Figure 6 Business Intelligence data and application location (Adopted from Gendron 2014)

Although cloud computing provides flexibility, cost savings and fast deployment times, it also includes downsides. Privacy and security concerns are common and service migration with other cloud providers or in-house systems might be challenging depending on the interfaces provided by cloud service. Moving data to the cloud can be more challenging than with internal servers. Many of the issues with cloud computing and BIaaS are related to lack of control of the environment. (Qian & al. 2009, Gurhar & Rathore 2013)

Data warehousing

Next we are examining data storing in more detail. A term data warehouse (DW) is often accompanied with BI solutions. Data warehouse is an independent data storage system with data that has been extracted from multiple sources to serve analytical purposes and application development. Inmon (1996) defines data warehouse as "subject orientated, integrated, non-volatile, and time variant collection of data in support of management decisions." The data can be collected both from internal and external systems. Data warehouses are built to support the decision making process and to serve as the foundation of a cross organizational BI system. Data models are built to support efficient querying, and the data is separate from that in operational systems. The operational data and data for analytical purposes are kept separate to enable smooth operation of operational systems. Running queries against databases is resource intensive and by dividing data both the operational and analytical processes can be optimized. The warehouse data is not directly linked to the data in the operational systems meaning that changes made in operational systems are not directly transferred to the data warehouse. The warehouse data is an offline copy of the operational system, data therefore needs to be refreshed by performing a reload on a regular pace. (Curtis & Cobham 2008, Kimball & al. 2008)

Storing data for analytical purposes is becoming increasingly multifaceted activity regarding both technical and logical aspects. In a world before internet era with limited choices of storing technologies, relational database technologies running on a company server were probably the solution for most needs. The data used for analytical purposes originated mostly from company's operational systems that also utilized relational databases. Traditional data warehouse deployments have been costly and taken minimum of six months before that DW can yield any business value (Kimball & al. 2008). Today's businesses change more rapidly than ever, vast amounts of structured and unstructured data are produced and companies that are able to utilize the new opportunities gain competitive advantages. The market offers new technologies and service models to respond to the changed environment. Therefore the traditional relational database running on company's server is no longer the only and most suitable option for all scenarios. (Dayal & al. 2009, Gendron 2014, Morton & al. 2014)

The type of data storage system depends heavily on the type and volume of data to be stored. If the data is less structured and contains various formats, documents, logs and pictures, it is necessary to consider NoSQL databases that perform better with different data types. If the data fulfills the criteria to be considered big data and therefore require special tools to cope with, distributed file systems and computing platforms need to be considered to achieve reasonable performance. Relational database is, however, still a valid and proven technology with structured data. (Gendron 2014, Morton & al. 2014)

In some cases a companywide implementation of a data warehouse may seem extensive. Maybe the benefits of a data warehouse are not so clear and the company wishes to start small and pilot a DW project through a departmental DW/BI project. It also might be that the data is not interconnected to a high degree and a centralized data system would not yield the benefits it should. A data mart is a smaller version of a data warehouse. It may be a part of a larger data warehouse system meaning that it is a subset of the whole warehouse data or it might be a standalone data storage that only serves specific business function. A distributed data storage system, however, introduces potential risks of data inconsistencies as the separate systems, and marts are difficult to manage. (Curtis & Cobham 2008) Kimball & al. 2008 dedicatedly leave the concept of data marts out of their Data Warehouse Lifecycle toolkit and present their view on data marts as a system that ends up mushrooming in to a data management nightmare.

In the next chapter we will discuss populating the selected data warehouse with business data.

Data analysis and modelling

Any company wide BI system usually pulls data out of multiple source systems. The source systems may be located in different servers and run on multiple platforms. Usually the initial business need is to get in touch with the data of individual source systems (Kimball & al. 2008). Source systems have a varying level of reporting tools of their own, and the reporting requirements for a centralized BI solution vary according to that level. In some cases the BI solution is not needed at all for core system reporting (e.g. cash flow) since the core system is capable of delivering the reports desired. (Kimball & al. 2008, Dayal & al. 2009)

The several sources of data represent a challenge for the vision of a cross-organizational BI platform. If operational systems are not integrated, they will most likely have different naming conventions and data types for same logical entities. The high level relations between source systems should be understood already when evaluating business cases and creating a project plan. In the data analysis and modelling phase the connections are identified in more detail and documented into the logical data model. (Moss & Atre, Kimball & al. 2008, Dayal & al. 2009) We will examine the data related activities by looking into a ETL process and its phases.

Extract, Transform and Load

Managing a data warehouse does not simply mean copying the data from one place to another. According to Alter 2002 the required tasks to maintain an operational DW system are the following:

1. **Extraction**, involves periodical loading of data from source systems,

2. **Consolidation/Conforming**, combining data from multiple sources to a logical data model,

3. **Filtering**, the process of eliminating data not needed for analysis purposes,

4. **Cleansing**, removing duplicate rows, correcting coding errors,

5. **Transformation**, modification of data so that it is consistent with DW data definitions, data formats and coding schemes,

6. **Aggregation**, the process of summarizing data to appropriate units for analysis,

7. **Updating**, loading fresh data periodically according to the rules defined for the data model.

Extraction is the first step of any ETL scenario. The raw data is extracted from multiple source systems to be processed to fit into the BI data model. During extraction it is necessary to be aware of distinct characteristics of the source systems that may have different database management systems (DBMS) and different connection methods. Extraction of data to the DW environment enables further processing of data independent of the original source systems. (Alter 2002, El-Sappagh & al. 2011)

Transformation of data consists of multiple activities. Consolidation or conforming of data unifies naming conventions for data that comes from different sources with different conventions. During this step same logical entities that are presented differently in different source systems are combined to a single entity. Another challenge with multiple source systems is columns that have the same name but they have different meanings. Some conformation challenges are presented in figure 7. (Inmon 2002)

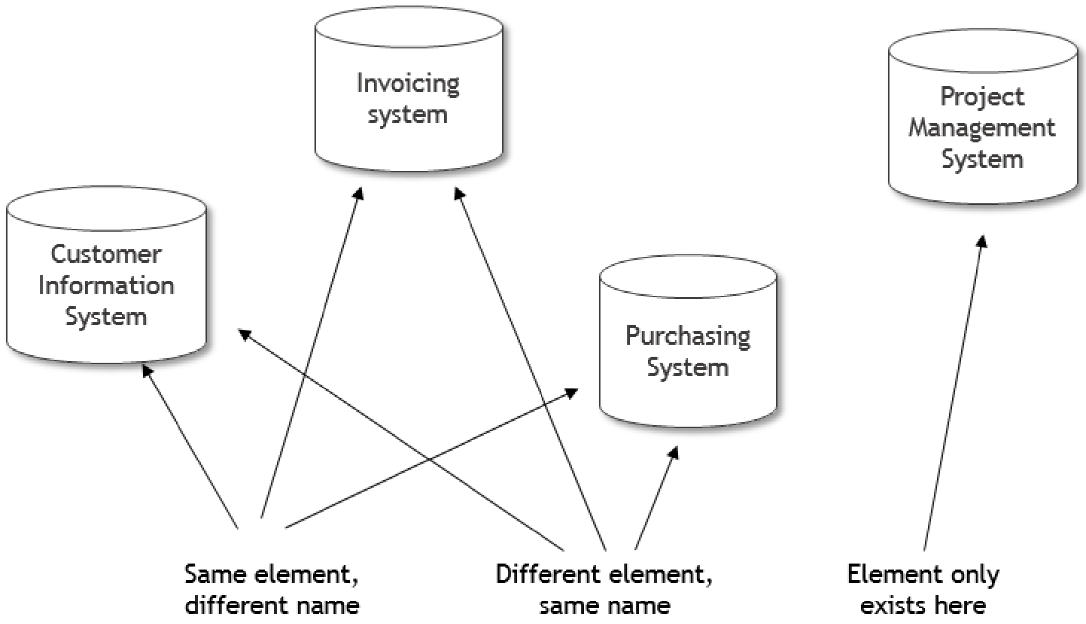

Figure 7 Conforming issues (Adapted from Inmon 2002)

An example of a conforming issue could be a "customer" field in sales system and customer relationship management (CRM) system. Both systems may have records of the same actual customer but they do not share a common key or id. During transformation process the two customer fields are conformed to a single field that serves the BI system. The data needs to be conformed also technically. Different source systems may present e.g. dates in different format. In order to be able to use all the dates in calculations in a same manner the dates must be transformed to a single format. This enables the application developers to do their calculations without having to worry about unexpected behavior of the data. (Alter 2002, Kimball & al. 2008, El-Sappagh & al. 2011)

Although it is important to have all required data in the DW, it is often necessary to leave something out to enable good performance and loading times of the system. Filtering source system log files and history data that is not meaningful for the current state of the business reduces the physical size of the data model and improves performance. Deciding what data to leave out from the system should not be only decided by IT staff since business users have the best understanding what is relevant and what is not. (Alter 2002)

"How much to cleanse?" is a question that needs to be answered in every BI project. Cleansing aims to improve accuracy of data by removing duplicate rows and correcting erroneous inputs generated by users. The data quality is determined to be one of the key critical success factors, and establishing good principles in data quality management within an organization is a cornerstone of an analytical culture (Dayal & al. 2009). The BI system is expected to offer an accurate representation of the state of business. It is also expected to show an accurate picture of the data that is captured by the operational systems. These two goals are contradictory to some extent and defining cleaning principles should balance between these two (Kimball & al. 2008).

The data quality problem originates from the daily business activities where operational systems are used in different ways. This leads to a situation where it is no longer possible to use certain fields of information as a base for analysis since they are understood and used differently in different parts of the organization. Another root cause for bad quality data can be low utilization rate of operational systems and use of alternative methods such as personal notes. In this scenario the analysis made on the basis of operational system data provides only a fraction of ongoing activities and may show just one side of the operations. Cleansing data when loading it to the BI database is not an activity that would improve the data quality within the organization in a long term. It only provides the polished view of the data and brings a larger portion of the data available for analysis. Any long term improvements originate from improved practices of using operational systems and are therefore responsibility of the business organization. It is important to be able to make the decision of not using poor quality data beyond cleansing for the analysis since it compromises the overall credibility of the system. (March & Hevner 2005). (Alter 2002, Kimball & al. 2008, Moss & Atre 2003)

External data

Different vendors are offering commercial data to be added to the company data repositories. This enables users to purchase data and accompany it with the internal data to find insights neither of the data sets could provide on their own. The needs and the availability of suitable external data depends heavily on the business and information needs. The external data needs to be integrated to the DW data model. (Curtis & Cobham 2008) The end users or application developers should not be able to identify the origins of the data by its behavior since it looks like any other DW data.

After the data is loaded to the designed logical data model it can be loaded to serve as the foundation for the next BI development steps.

Application development

An interactive report that a user uses for viewing and analyzing information is called an application. An application may e.g. consist of a monthly sales report with possibility to filter data or it might be aimed for company executives to offer a holistic view on company performance. This chapter describes different types of applications, related concepts and technologies.

Standard reports

A modern BI platform should offer different level of detail and freedom to different types of users. The most structured and predefined level is standardized reporting. A company needs to monitor past and current performance to be able to determine whether they are doing well enough and meet the set target levels. Standard reports consist of selected metrics often called key performance indicators. A benefit of a BI system regarding performance reporting is the capability to automate the reports so that they do not require any periodical work of employees. Data gathering and calculation of metrics occur as background processes. Automated reporting also is less prone to errors since the metrics, data sources and filters are fixed and not touched by anyone between reporting periods. The focus on standard reports is solely on accuracy and accountability with minimal human intervention that might unintentionally alter the results. Figure 8 illustrates a standard report mockup. Standard reports can be shared through an information portal or sent automatically with e-mail to defined recipients. (Loshin 2013, Kimball & al 2008)

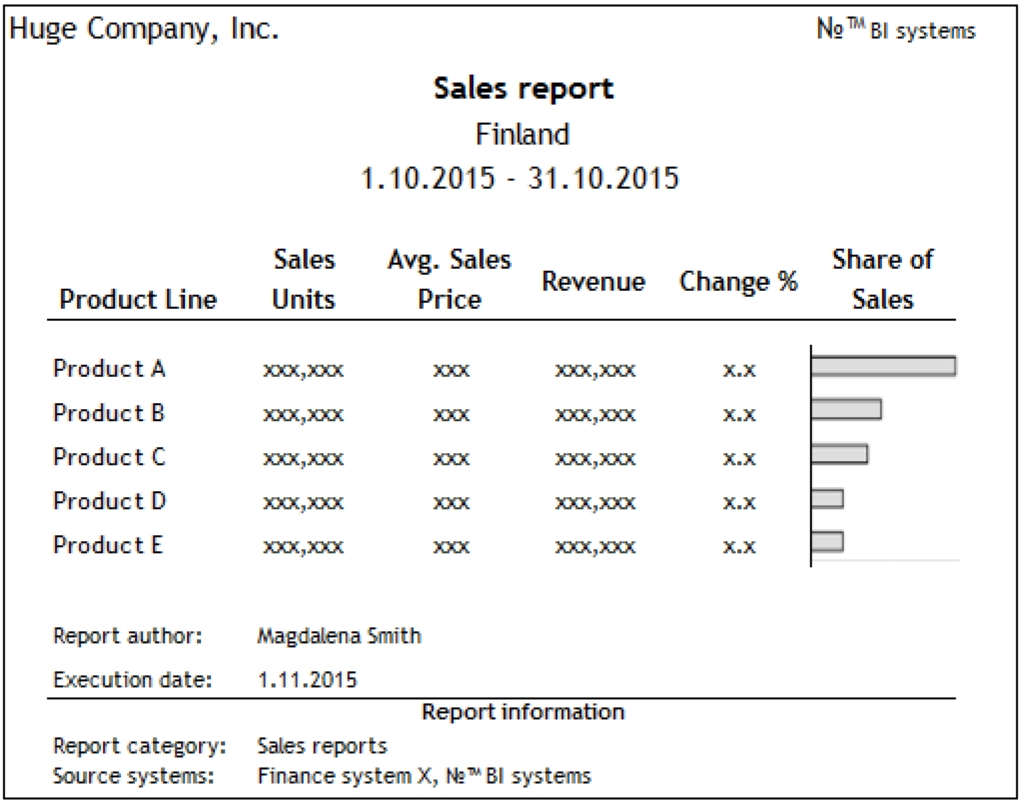

Figure 8 Sample report mockup

Data discovery

The next level is to offer more flexibility and different angles to the high level key performance indicators. The aim is to provide users with the possibility to adjust the level of detail and abstraction. A well designed interactive report enables users to delve deeper into the high level metrics when they wish to see the detailed events that add up to departmental or areal figures. The application must offer a quick-to-grasp overall picture and various level of detail and different angles to the information. Different users require different level of detail. Managers may be too busy to find reasons behind results but are very interested in quick access to KPIs and delegate the reason analysis to their employees. (Yi & al. 2007) When offering high level KPI figures to manager it is of paramount importance to be able to answer questions that arise from the metrics. Presume that company KPI A shows value of 112% and the responsible manager tells you this cannot be possible. Without the ability to look into details and show the events behind the KPI it is easy to overlook the results provided by the BI system if they do not match gut feel. The credibility of the BI system is highly reliant on being able to easily validate the results and provide visibility to row-level data. (Kimball & al 2008) This can be achieved with more sophisticated applications that offer a graphical user interface that allow representing information with collection of graphs, tables, pivot tables and lists. Applications can offer filtering options and drill-down possibilities that enable deeper understanding about the KPI results and comparison between geographical areas and departments. Drilling-down means the ability to increase granularity of the portrayed data. Basically this means that the user is able to see e.g. individual rows that add up to a higher level results. Depending on the application development tools in use, the drill-down capabilities need to be created into the underlying data model during the ETL process. This means defining hierarchies e.g. that cities are under states or that months are under certain quarters of the year. (Kimball & al. 2008, Curtis & Cobham 2008)

Dashboards

Overall views on company performance are called dashboards. The term originally meant the piece of wood that protected the driver of horse driven carriage from splashes and dirt. Later it was adopted by automotive industry where it served as the console that separated the driver from engine room and provided information about the functioning of the vehicle. Cars utilize the concept of dashboard in a very similar way to business analytics. The most important key performance indicators are presented in a way that enables quick overall comprehension about the current state of the car. Whereas the dashboard of a car shows speed, fuel levels, travelled distance and warnings in case of malfunction, the company dashboards represent e.g. amount of sales, profit margins, backlog and work in progress (WIP). (Curtis & Cobham 2008) Figure 9 shows and example of a dashboard.

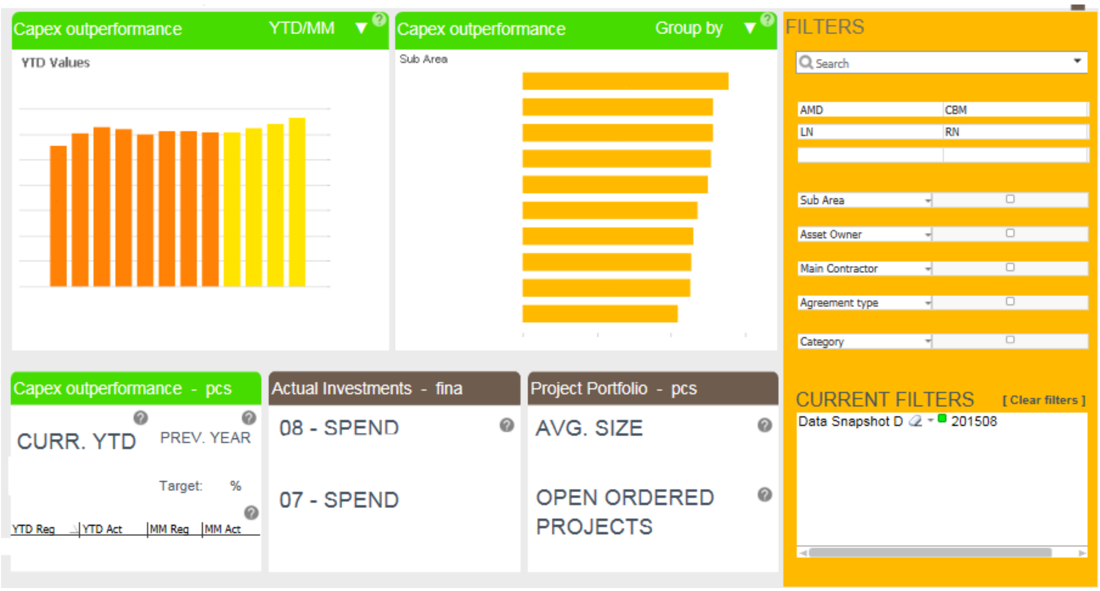

Figure 9 Caruna asset management dashboard (numbers hidden)

Information visualization

Information visualization plays a high role in application design. The significance of colors for human perception is a long known fact identified by cognitive psychology. According to Carter 1982 people are able to find information significantly faster if it differs in color from its surroundings. This is often utilized in BI applications that color results according to defined target levels. This speeds up the use of portrayed information. If an application highlights deviances from target levels e.g. with red color, a person viewing the application does not have to really look into single values if he or she is only interested in seeing whether metrics are above target levels. Coloring schemes that are familiar to people from general contexts enable users to intuitively understand which colors are for good performance and which for bad. An example of a general context color coding is the red-yellow-green color coding used in traffic lights. It is no surprise that traffic lights and their coloring is a widely used concept in BI information visualization. (Moss & Atre 2003, Curtis & Cobham 2008)

Data mining

Data mining is the most challenging level of BI. It might be difficult to design data mining applications since data mining often occurs adhoc and involves different data for each analysis. Therefore data mining often requires an environment that is directly aimed for data mining activities. (Moss & Atre 2003) Data mining techniques are used to find patters and correlations from large datasets (Alter 2002). It involves techniques such as multiple linear regression where a number of variables is examined to identify patterns. Data mining techniques provide answer to more open questions compared to reports and dashboards. Rather than direct questions with definitive answers such as 'In which areas we sell the most?', data mining techniques seek to identify patterns that possibly underlie under more easily observable facts. By finding distinctive characteristics that might correlate with high sales from the top selling area, ways of operations or for example legislation, the company has the ability to pursue those characteristics on other areas as well. Data mining techniques include (Bocij & al. 2008):

1. **Identifying associations**, a supermarket purchase data might reveal that customers that bought lightbulbs bought ladders more often than other customers. It is not always however clear how to use information of associations in a meaningful way,

2. **Identifying sequences**, by examining sensor data from production facilities, it can be possible to identify signals that precede failure. A company can create a predictive maintenance program that reacts to the identified signals,

3. **Classifications**, identifying customer groups that share similar buying habits. A more targeted marketing campaign can be created,

4. **Clustering**, finding groups of facts that were unknown before,

5. **Forecasting**, using yearly sales data to determine production levels.

Deployment and governance

A BI system deployment should follow the same iterative process that is present in the whole BI initiative. The incremental roll-out reduces the risk of exposing the whole organization to the possible defects and difficulties that the initial roll-out often involves. Introducing BI applications to a small group of people with varying roles and technological skills enables the development team to identify usability issues and defects in cooperation with the users. A well-organized roll-out phase plays a big role on how people adopt new tools and how they feel about them. Training and ongoing support will help to gain the user buy-in. (Moss & Atre 2003)

Metadata

To assist users new to the BI system the solutions should be documented in a way that helps to understand the system. Metadata is often described as "data about data." (Inmon 2002, Gardner 1998, Foshay & al. 2007) The purpose of metadata is to help users and developers to better understand contents and rules of a BI system. Gardner 2008 describes metadata as a map that enables users to navigate in a data warehouse environment and help them to find what they are looking for. He compares BI system users to library users looking for information and metadata as a library catalog. Library catalog lists available books and their location and provides information such as publishing year, author and possibly related works. Library catalog enables users to get the most out of library collection. In a similar way a BI environment can be used more efficiently if users can easily find the reports and applications they are looking for.

A companywide BI environment offers information from multiple functions, systems and authors. Metadata helps users to place information to context and provides guidance for finding related phenomena and causes behind metrics (Inmon 2002). Different categories of metadata try to answer different types of questions. Technical metadata serves the BI environment developers helping them to maintain and develop the environment. Business metadata helps users in using the information located in BI system. Metadata should be made easily accessible for the users so that they actually want to use it. (Moss & Atre 20003) Foshay & al. (2007) categorize metadata on four different categories presented in table 1:

Table 1 Metadata categories

Category	Explanation
Definitional	Definitional metadata answers the question: What does this data mean, from a business perspective? This category includes business definitions, allowable values, calculation and business rules.
Data Quality	Data quality metadata answers the question: Is this data high quality enough for me to be used in my analysis? This category informs users about the freshness, accuracy, validity and completeness of the data in the system.
Navigational	Navigational metadata answers the question: Where can I find the data I need? Navigational metadata provides users means to find the data they are looking for and provides information about relations of the data.
Lineage	Lineage metadata answers the question: Where did this data originate (source systems and files) and what's been done to it (ETL steps and transformations)?

Developing analytical culture

The deployment of a BI system is a great opportunity to lay the foundation for improved analytical culture within the organization. Olszak & Ziemba 2007 state that any long term success of a BI system relies on continuous improvement initiated by both technology and business users of the system. Businesses change rapidly, which requires continuous update of data models in use and incorporating new data sets to the company data model. To fully utilize the company data model, the users should be provided with tools so that they can create their own reports in cooperation with BI development team.

The culture change does not only include developing new applications on an ongoing basis but the way people look at information needs to evolve as well. Interpreting the results in an environment that provides freedom and flexibility requires skill as well, and people need to learn to ask sophisticated questions to further understand the information. Ultimately, the increased availability of information should improve business and decision making. This is not self-evident since connecting information derived from data with real life situations may prove challenging.

Enabling smooth operation

After the system is taken in to production it is necessary to establish maintenance principles for the new system. The data models, applications and reload procedures are likely to require changes along with software updates. The system needs to be able to recover from unintentional changes that occur during the development activities. A backup of the data and applications should be created on regular intervals so that in case of failure it is possible to go back to the situation that is close enough to the present time. This minimizes the amount of rework that is required to resume to the situation before error. In a best case scenario all the definitions and code are up to date and the only thing required is to load the newest data. (Kimball & al 2008) Backing up the data is especially important when dealing with very large datasets. Reloading all data from source systems after system failure is very time consuming and takes down the BI environment for long periods of time if the amount of data is large enough. (Moss & Atre 2003)

Business intelligence success measuring

It is widely accepted in the literature that the use of business intelligence systems can bring multiple benefits and increase the profitability of the business with one way of another. Due to the nature of BI systems the actual benefits are often difficult to weight on accurate terms. Lönnqvist & Pirttimäki (2006) conclude that the BI measurements serve two purposes. The first and most common reason is to determine whether or not the BI system is worth the investment. A traditional way to asses a profitability of an investment of any kind often includes return on investment (ROI), net present value (NPV) and internal rate of return (IRR) calculations. They can be used also to assess success of a BI system. (Saarinen 1996, Hocevar & Jaklic 2010). Gessner and Volonino (2005) assess success of two BI initiatives related to customer transactions tracking and management through return on investment. They conclude that ROI was possible to calculate for transaction tracking initiatives and show that they were highly profitable. It is, however, a rare case when direct monetary evaluation is sufficient or even possible. Saarinen (1996) argues that the outcome of an information system project is not actually the system itself but the benefits of utilizing it for improved business decision making and actions. Assessing how a system serves users in their daily work cannot be concluded with exact terms so qualitative characteristics of the system need to be considered along with raw financial calculations. Qualitative evaluations help to serve the second purpose of BI measuring. The qualitative findings help to manage and develop BI processes by adopting best practices from successful projects and avoiding performing the same mistakes on a continual basis when they are identified. Some benefits that BI systems deliver are highly intangible, such as improved control over processes, improved communication within the organization and higher quality of available information. The benefits should eventually lead to improved financial performance but there is often a time lag between the production of the information and financial gain. (Lönnqvist & Pirttimäki 2006)

Saarinen (1996) has developed a widely recognized information system evaluation model that incorporates both financial and intangible benefits assessments. Although it is a model aimed at assessing information systems in general, it provides reasonable framework for BI success evaluation as well. His model consists of four main dimensions. The dimensions of success are: development process, use process, quality of the IS product and impact of the IS on the organization. The model of Saarinen is an extension of User Information Satisfaction (UIS) evaluation model. The four dimensional model tries to address the shortfalls of original UIS by incorporating evaluation of development process and the quality of the IS. UIS is used to measure the attitudes of users towards the information system. Ives & al. (1983) conclude that the idea behind UIS is that an information system which meets the needs of its user will reinforce satisfaction towards system. On the other hand, if the user is dissatisfied with the system, he or she will look elsewhere and try to use alternative methods. The satisfaction and level of devotedness to the systems reinforces the benefits that can be gained by using the system and is therefore a significant factor when evaluating system success. The development process dimension assesses if the development process was able to provide solutions that respond to user requirements within time and budget constraints. If the development process consumes more time and resources than reasonable for the outcome, the financial outcomes of the project are not likely to be successful. The use process dimension assesses if the development team has been able to capture the user needs and respond to them with the IS solution. System quality encompasses both system and information quality of the system. System quality refers to the usability and responsiveness of the system whereas quality of information refers to timeliness, accuracy and completeness of the provided information. The four dimensional model of Saarinen is presented in figure 10.

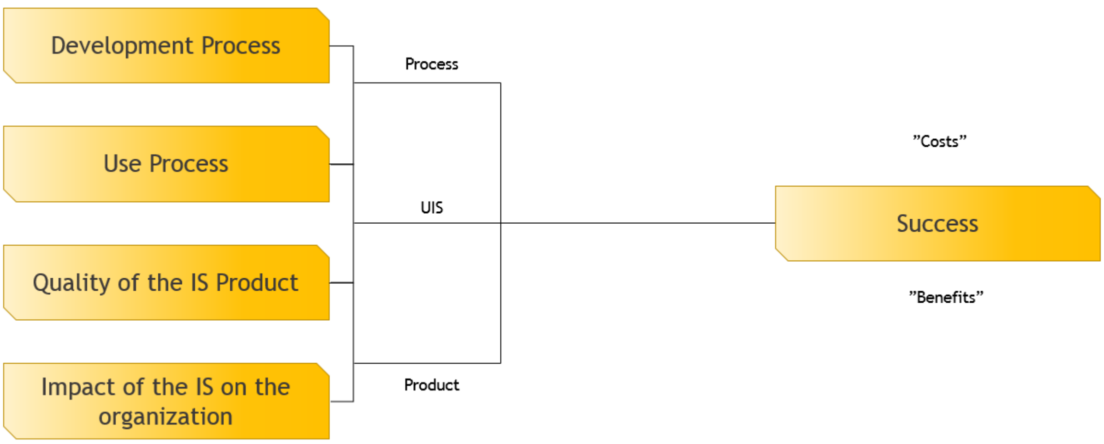

Figure 10 Main dimensions of information system success (Adopted from Saarinen 1996)

Each BI solution is targeted for individual purposes and the criteria to assess whether or not they were successful are unique in each case as well. Any assessment model can therefore only point out where to look.

Chapter summary

This chapter provided a literature view on the matter of business intelligence deployment. The deployment framework was presented in a chronological order from initial financial and need assessments to post deployment evaluation of the BI solution. The pre-development steps in the presented business intelligence deployment framework were evaluation of business cases and needs, evaluation of organizational readiness for a BI system through success factors, gathering business requirements and project planning. The development phases consisted of technical architecture choices, data analysis and modelling and application development. The deployment and governance along with measurement of BI initiative success were the last steps in the presented business intelligence deployment framework. In the next chapter we will evaluate two different business intelligence report development projects regarding the same system with different tools and different technical and organizational environments. The literature review and the presented deployment framework was provided to be able to better assess and understand the case study research in the next chapter.

4 Case study research

The project control system report development case serves as the main instrument of evaluating the differences between Fortum era and Caruna era report development principles. Since the same functionalities from same source system were developed by Fortum corporate IT and later by Caruna BI staff using distinctively different methods, this case serves as the most thorough single initiative for method and result evaluations. Both solutions were also relocated to different servers after less than half a year of use. The relocation of reporting environments with traditional DW and with newer file-based technology proved to be far apart from each other. The case evaluation presents the development initiatives first and relocation projects after that. The Fortum era solution is evaluated first and then Caruna era solution is compared to it.

Caruna invests approximately 100M€ yearly in network assets. Investments consist of low and mid voltage overhead and underground lines, transformers, electricity stations and other network components. All building activities are carried out by subcontractors. The work is performed under different purchasing contracts which apply for single orders or longer term partnerships with agreed volumes and prices. Roughly a year before the divestment of Caruna a three-year partnership contracts were made with the largest contractors which defined prices, yearly volumes and delivery terms for grid building projects. The contracts were called Radisson contracts. To monitor the vendor contract compliance, a Radisson reporting project was started. The goal was to identify how well contractors are able to meet agreed milestones, compare vendors to each other and to determine financial efficiency of different contractors.

Fortum era Project Control System reporting

The operational system for grid building projects was called PCS (Project Control System) which was used by all Nordic organizations. Corporate IT provided two data warehouses for distribution organization. One was administered in Finland and other in Sweden. The project management organization was on a Nordic level and led from Sweden. The reporting solution was to be created on Swedish DW that already had PCS data modelled and ready for use. The definitions for reports were gathered both from Finland and Sweden since the countries used PCS in slightly different manners. The main responsibility for report development was on an external consultant who was expert on the reporting software in use. The report loading scripts were developed by BI team manager in Sweden.

The first versions of the Radisson reports were developed rather quickly by the Swedish BI team. The reports, however, were easily identified as not working correctly. The reports only included a portion of data and visualization of the reports was on a very low level. The author returned to the company as an employee after a half a year of study break and started troubleshooting and developing the reports to a more complete level. The author was not dedicatedly hired to work in the project but started to work on it mostly due to my previous experience of the PCS reporting before the study break. At the time the author joined the project, it seemed that no further involvement of business people was planned and any testing related matters were left unattended. The organization had participated in business requirements gathering and provided documents to the IT team that were supposed to provide all required information regarding the business needs. The documents, however, were incomplete and contained some mistakes. In addition to incomplete specifications, the IT team had interpreted the specifications slightly differently than they were originally meant to be understood. The collaboration between business representatives and IT team was lacking means to be in touch on a daily basis in order to clarify ambiguous definition documents or business needs. This was partly a result of people working both in Finland and Sweden. A more significant factor, however, was the mindset that the business organization showed towards the project. The business ordered the BI solution from an external party and expected to receive a working solution in return. No real collaboration was planned, and therefore the project was hanging in the air after the IT team had responded to the definitions as well as they could. The initial setting of the project failed to meet the project managerial and team related success factors discussed earlier in the literature review.

In order to meet its goals the project needed someone that had time and interest to interpret the project results against the business requirements and rules. The managers were too busy to actually look into the results and try to identify reasons behind false figures. The error tracking process was started by trying to create metrics defined in requirements definition documents by hand using spreadsheets and data export that the source system PCS was able to produce. Finding what was wrong in the automated reports did not require significant amount of work and correcting steps were quite easy to identify.

The next challenge was that although it was now clear what needed to be done in order to fix the reports, the technical control over the solution was strictly in the hands of IT staff. Every change had to be explained to a person who had very limited understanding of the business and content and characteristics of the source system. When a defect was identified and communicated to the IT staff, it had to be concluded whether it originated from the loading script or from the application layer that were developed by two different persons. It seemed that the BI team manager had very little time for the project after he had written the initial loading script. Making changes to the script required weeks of waiting before the changes were passed to the application development phase. The application developer made changes and passed the results for business for inspection. Often correcting one mistake revealed others that were hidden by the original mistake. Identifying, fixing and inspecting of errors had to be started from the beginning. The development speed was painstakingly low when a simple iteration cycle that was able to correct few lines of code took from two to four weeks of calendar time.

At this point the business users started to feel that it requires more time from IT staff to understand the business problems than it would require from business people to understand technical solutions. The IT staff had traditionally been very protective over their systems and DW data. One reason for this was that Fortum was operating in both electricity sales and distribution businesses that are separated by law in Finland. Distribution companies are natural monopolies that are required to serve sales companies equally without sharing knowledge to any parties that might gain competitive advantages from the information. The data warehouses supported both sales and distribution businesses, which amplified the protective culture the IT management showed towards user privileges and information openness.

Gaining read access to the DW to see what was possibly wrong required some pressure from the business managers. In addition to this, user rights to reporting platform were granted. When the report developer and the author as a tester got access to the same script and tools, the communication accuracy improved significantly. It was no longer required to communicate application defects by simply telling that "We are receiving less data than we anticipate." From this moment the tester had the ability to examine loading script and the applications to identify possible causes for the defects. Although the tester did not fully comprehend what he saw, he understood enough to guide the developer to look into the right parts of the solution. When the developer explained how specific details in the system work and relate to other components, the tester and developer were able to trust that they had the same understanding of the matter. Problems were solved faster and with higher confidence than before.

Evaluation of Fortum era Project Control System reporting project

The Radisson reports were taken into production relatively quickly after the tester and developer gained visibility to the same tools and information regarding the technical solution of the reports. The key enabler in this project was crossing over strictly defined roles and responsibilities regarding the project. The initial project roles were too limited and did not enable the people to look into the work of the others and mutually explore different solutions. People and project areas acted as black boxes that received input from other black boxes and produced their output with very limited understanding about anything else than what they directly perceived. The difficulties in the project were mostly result of poor project management, too strictly defined responsibilities and lack of communication. These factors built on top of each other and were amplified because people were located in different countries. The responsible business manager was also replaced during the development project. The BI manager, report developer and business manager did not meet during the project, and nearly all communication between business and IT was routed through the tester.

Caruna era Project Control System reporting

The first steps in building Caruna’s own analytical platform took place approximately six months after the Radisson reports were taken into production. Relocation of PCS reporting from Swedish to Finnish DW had been going on for nearly four months which meant that the systems were offline from business perspective. Frustration towards the lack of control over business information needs and solutions was growing and development of anything new was considered extremely difficult.

By that time, a new business controller joined Caruna. He had previous experiences of a more agile BI solution that did not require development of data warehouse, was very cost effective compared to a DW and reporting environment implementation and had received very good feedback in his previous organization. After realizing the current situation and future needs for being able to operate without corporate IT as a standalone company, the business controller initiated an investment project for acquiring the more agile technology. Since the initiative for new analytical platform came from business manager, the development team received one of the most fundamental critical success factors from the start – committed managerial support.

The most determining factor for the new software acquisition was to gain better control over information assets. The lack of control was not in the same level at all parts of the organizations. Some business systems provided better access and visibility to the data. The people who had previously been working with Finnish BI team had closer relationships to BI developers. Therefore it is not surprising that new BI platform development efforts began in the part of the company that would benefit the most from improved control. The system manager of the PCS system was also appointed to the role of new BI platform system manager. His role was to keep servers running, to ensure that hardware and software components align and that connections from business systems to BI servers can be established. My role in the new BI platform implementation was to familiarize myself with the new technology and to develop the first data models and reports in cooperation with external consultants.

The setting for PCS reporting development at this point seemed rather good. Knowledge of the business processes was familiar at some level both BI system manager and application developer. Both knew the organization well and were able walk next to business users to ask questions. Access and understanding of PCS source system was provided by BI system manager. The question at this point was whether or not this setting could be utilized by using the new technology to actually deliver accurate and consistent reports.

The first PCS reporting applications that utilized modelled data layer were introduced after 3 to 4 months after the software acquisition decision was made. The first applications provided increased visibility to the state of network building projects and were the first attempts to automate KPIs of asset management. Some metrics required from 2 to 3 months of more work before they were trusted. This was, however, the first time that the most challenging KPIs were used from the BI platform instead of analysts building them each month using spreadsheets. It was significant as well that quick fixes to calculation rules and to the data model did not any more require time from 2-4 weeks. The easiest problems were solved in minutes since people were sitting close to each other and had the means to influence the platform.

This was something quite new for an organization that had waited approximately 6 months for a report that already had the data layer in place. In addition to organizational and application development project related factors, the new technology also had some features that aided in achieving the level of performance. The new software was very quick to learn for someone that understood relational databases and SQL. Kimball & al. 2008 describe multiple sub-processes and phases that take place during an ETL development phase. Many of these seem extensive from a viewpoint of a company of 300 employees that operate in a single country. Instead of building a database optimized for querying, the new technology enabled storing of data as files. Data was extracted from PCS and transformed using single script that stored the table files to the server. Files appeared in file manager as any other type of files and could be moved and copied in the same way as photos or text files. The data files were optimized for the BI tool which provided high performance in terms of query speed. This was beneficial during application development since playing with data was fast and convenient. The applications loaded the data into RAM for fast processing.

Evaluation of Caruna era Project Control System reporting project

The main reason for the success of Caruna era PCS reporting project was the fact that the organization had a team of two persons with control and understanding over the whole development process. The network building project business was familiar to both, and translating the user requirements to technical definitions was easier than in cases where technical staff only understand the technical aspects of the solution. Since PCS data model and reports were first outcomes for the new BI platform, the design choices could be made to support this specific implementation with no need to fit into a previously defined structure. The new BI software supported this approach very well and enabled fast iterations of reports and data model.

Business intelligence platform relocation

Fortum’s distribution organization was prepared well in advance regarding the future separation. The distribution businesses were sold one country organization at a time. Since the organization had been on a Nordic level, some organizational changes were performed. This was arranged so that the companies offered for divestment would have their own organizational structures and the ability to eventually transform into a standalone company or to merge into a larger organizations individually. Part of this process was the relocation of asset management BI solutions that were hosted by the Swedish DW team. The Finnish DW team had previously served network operations and customer services, and this was the logical point to bring asset management organization also under same BI services. The relocation of the BI solution required establishing DW data models and ETL processes and transferring the developed report applications to the Finnish platform. The relocation project stretched over six months of calendar time. Out of that six months the reports were un-usable for four months. The Finnish data was removed from Swedish DW after the initial relocation, and the Finnish environment experienced difficulties in getting the system error free, up and running.

The effect of new difficulties in the business organization was very demoralizing. The Radisson contracts had been in use for one year and out of twelve months the reporting capabilities were operational from three to four months. After the relocation was close to complete the business was already looking for new types of contracts, and the need for Radisson reports decayed quickly.

The same reasons that were behind the difficult development process of Radisson reports were also present in the relocation project. It seemed that the project was not properly resourced and the overall responsibility over relocation of PCS related reporting was vague. The Finnish DW team was considerably more demanding regarding the communication channels related to the relocation project. The team insisted that users use a ticketing system, where users provided request through the system that directed the ticket to the right person. Requiring business users to communicate through tickets on an ongoing relocation project did not encourage cooperation and mutual responsibility over the success of relocation project.

BI platform relocation with file based software

New Caruna BI environment was originally established in to a Fortum network where servers and server management were provided by a partner. Eventually the BI environment had to be moved to Caruna's own environment and servers as part of establishing Caruna's own IT infrastructure. Caruna also chose to partner with an IT service provider which was responsible e.g. for hardware, servers and networks. The relocation task was in terms of overall complexity less demanding compared to Radisson report and PCS reporting relocation in the old environment. The PCS related data models and reports had to be cut off from the original DW/BI environment and then integrated to another DW/BI environment. In this later scenario the whole system was moved from one service provider's server to another service provider's server. Therefore the latter case included less integrations and less decisions that needed to be answered.

The relocation of a file based BI environment proved to be extremely straight forward. The software was installed to the server, required firewall ports were opened and database connections established. All the existing data, data models, applications and loading scripts were transferred using simple copy-paste commands. The execution and timing of scripts had to be configured manually which was half a day's work in an environment consisting of four source systems. Most reports and applications did not suffer any downtime, and the relocation took two weeks of calendar time from start to finish. One source system data connection, however, could not be established in the new environment. The reason behind this was that this network information system (NIS) was running on a server that required opening of four different firewall ports. This was not foreseen in advance but fixing the problem should have been very easy. For some reason, however, the ports had not been opened even after 2 months of waiting and repeated inquiries. The same reasons that prolonged BI development during Fortum time started to influence development and administration work also in the environment that should be in control of Caruna. The cooperation with the service provider started to show signs that might cause problems in the future regarding BI development work. The main cause of problems were difficulties in connecting different source system and servers due to very strict firewall policies that the service provider enforced. This problem was not strictly BI related but influenced also other IT related development work. The actual downtime of NIS data source link was closing to three months during the time of writing this work. Identifying the reason behind connection did not begin until after one month of downtime since NIS reporting was less critical than other areas of reporting.

Overall evaluation

Fortum and Caruna era solutions were developed within very different organizational situations and with different technologies and processes. Comparing the two different BI teams or the two different technologies in terms of proficiency to deliver would be unfair. A more fruitful approach for finding best practices for future development is to identify the effects of development process factors to the actual end result. The debate over best BI management and process principles can be continued perpetually if it is not measured by the ability to answer information needs of business users. Inadequate metrics for measuring the success of BI may actually harm the actual mission of serving business users. The asset management organization experienced this during Fortum corporate IT BI development times. The DW team had the ability to meet service level agreements (SLA) as a key bonus metric. This encouraged communication through ticketing system which recorded data about this ability. The BI related development problems, however, required face-to-face meetings in order to achieve mutual understanding over the matter. Requiring the user to mainly use tickets as the primary means of communication hindered the speed of development and time for resolution.

The Fortum era and Caruna era PCS reporting solutions are like day and night if evaluated side by side. The Caruna solution achieved quick adaptation from business users and was able to respond to emerging needs very quickly. The used ETL logic, however, was not scalable for larger source systems and the documentation was very limited. The Fortum era BI solution and management principles were very well defined and followed but the ability to deliver was significantly worse.

Chapter summary

The BI deployment case study research in this chapter evaluated two report implementation projects regarding the same source system. The projects utilized different tools, development methods and technical architectures. The evaluation was made by comparing the Project Control System report development projects of an electricity distribution company Caruna before and after it was divested from Fortum. The successes and failures correlated with the success factors proposed by literature, and the organizational factors were identified as the most influential. The case study research and literature review are used together in the next chapter when providing conclusion and recommendations and managerial implications.

5 Conclusions and recommendations

The importance of the ability to use information for competitive advantage is increasing rapidly. The DW/BI principles promoted by Kimball & al. 2008, Inmon 2002 and Moss & Atre 2003 that have served as a cornerstone for BI development do not anymore suffice in modern BI field characterized by big data and self-service concepts. The concepts are still applicable but the organizations are required to understand the changing nature of information and available tools to be able to produce information cost effectively and fast enough. The literature review did not identify any widely referenced business intelligence book that would cover the current BI area as a whole and provide a holistic picture of current state as the three books mentioned above once did. It is important to understand that data governance requirements are very different for multinational companies compared to a company operating in a single country. The literature often suggests heavy governance models that take into account nearly any imaginable scenario. From a standpoint of a national company with 300 employees and somewhat coherent information systems, following governance models by the book would slow down development work and increase costs without delivering benefits of same scale. The information governance overhead should be proportional to the complexity and size of the organization and its information assets.

The case study identified the cross organizational knowledge, overlapping roles and low technical proficiency requirements of the BI platform as the key components of the improved ability to create BI solutions for business needs. Factors that decreased productivity and contributed negatively to overall ability to create, deliver and maintain information to business users were mostly related to the organization rather than to technology. Strictly defined roles and overly tight access control of users and server connections created bottlenecks that prohibited competent people from refining data into information.

The question for Caruna during the time of writing of this work was whether or not they could achieve same results as with the first reporting initiative throughout the whole organization in the changing business application landscape. The setting where from two to three people have nearly all the required information for delivering a BI solution is not common within an organization by default. A key component of future

Caruna BI strategy will be identifying people who work in the organization who possess business understanding and have an analytical mindset. By providing the BI tool and access to data with support of the IT team the aim is to create analytics in close collaboration with the business. This ideology takes the standpoint that it is more beneficial in a long run to educate the business people in new technologies and move the responsibility of the content of the reports and applications to them. The IT works as an enabler rather than as an information vendor who fulfills request upon order. The information exchange is expected to work both ways and eventually lead to a situation where information is in the heart of making decisions and running operations. It is, however, required for Caruna to establish framework for BI development to ensure control over growing data model and number of information workers and consumers.

Managerial implications

From managerial perspective the agile development methods regarding business analytics require fresh perspective to the information governance. Self-service concepts that enable larger portion of the organization to create and understand valuable business information cannot be managed with strict hierarchical policies. It is, however, necessary to retain sufficient level of control to prevent the system from collapsing to its complexity.

An independent research company Gartner presented three key dilemmas in their business intelligence summit in München during fall 2015 that characterize the managerial decisions related to business intelligence. The first dilemma was centralized vs. decentralized BI management. The presented dilemma underlined that decentralized development of reports and applications does not produce coherent and unified solutions in a way that a centrally managed BI function does. Decentralized methods possibly produce information faster without IT involvement but it may prove difficult to combine the produced data models and solutions with other data and information.

The second dilemma was certainty vs. uncertainty. This dilemma pointed out that the same rules apply to business intelligence management as in any other field of business.

Old proven technologies utilizing company internal data provides high level of certainty regarding costs and benefits. The performance of such a system however is easily outperformed by a successful BI system utilizing emerging technologies' self-service concepts. The latter, however, introduces more uncertainty regarding information consistency and correctness along with actual costs of implementation and operation.

The third dilemma was share vs. protect. The dilemma presented that the transformation of the BI field is likely to cause heavy change resistance in organizations that are used to centrally managed business intelligence. With new technologies, more and more users are able to create their own solutions and answer questions without the help of IT department. Gartner presented that denying access to data and information often serves the purpose of individuals rather than the good of the company. By restricting information, knowledge and tools from others, individuals are able to promote their own status and position in the company. All data, however, cannot be shared. Personal information of employees and customers must be protected against unauthorized use.

A successful business intelligence strategy is able to balance with the two ends of the three dilemmas. The emphasis should be in controlling the factors that matter the most. In centralized vs. decentralized dilemma the control should be in governing the data models and technical architecture to support less controlled decentralized BI development. Certainty should be pursued in most vital information that is used in decision making on a continual basis. Company should create a BI platform and culture that is able to quickly respond to new information needs and data sources to cover uncertainty. Companies need to move away from restricting user rights by default and assessing whether or not an individual should have access to information. Sharing should be the new default, and access limitations are reasoned and explained. The control aspect protects information from getting into wrong hands.

The themes highlighted by Gartner were strongly present in the case study research. It cannot be stressed enough how large role the management principles play in the successes and failures of BI initiatives.

Recommendations for further research

There seems to be a lack of a comprehensive business intelligence book that would cover all areas of modern BI field and present the most important milestones with related decisions regarding BI system implementation and development. The books that have held this place such as Inmon 2002, Moss & Atre 2003 and Kimball & al. 2008 do not keep in pace with the emergence of big data, self-service and new platform and database technologies. Big data related publications are available in form of books and scientific articles. The business intelligence self-service area is poorly covered by literature, and businesses and individuals have to resort to less scientific publications, surveys and marketing speeches. Currently the most up to date and trustworthy information regarding business intelligence field is probably provided by Gartner.

6 References

Alaskar, T & Efthimios, P 2015 June, 'Business Intelligence Capabilites and Implementation Strategies', *International Journal of Global Business*, 8 (1), pp. 34-45

Alter, S. 2002, Information Systems: The foundation of E-Business fourth edition, New Jersey: Pearson Education

Atre, S 2003, 'The Top 10 Critical Challenges for BI Success.' *Computer world,* special advertising supplement. url: http://www.atre.com/pdf/BI_top_101.pdf [7.5.2015]

Carter, R.C 1982, 'Visual search with color', Journal of Experimental Psychology: Human Perception and Performance, Volume 8, Issue 1, February 1982, pp. 127-136

Chaffey, D. & White, G. 2011, Business Information Management second edition, Essex: Pearson Education

Chang, V. 2014, 'The business intelligence as a service in the cloud', *Future Generation Computer Systems*, *37*, 512-534.

Curtis, G. & Cobham, D. 2008, Business Information Systems, Analysis, Design and Practice 6th edition, Essex: Prentice Hall

Dayal, U., Castellanos, M., Simitsis, Alkis. & Wilkinson, K. 2009, 'Data integration flows for business intelligence', EDBT '09 Proceedings of the 12th International Conference on Extending Database Technology, pp. 1-11, New York, USA

Elbashir, M. & Williams, S. 2007, ‘BI Impact: The Assimilation of Business Intelligence into Core Business Processes, *Business Intelligence Journal*, (Fourth Quarter 2007), pp.45-54 (ladattu)

El-Sappagh, S.H.A., Hendawi, A.M.A. & Bestawissy, A.H.E. 2011, 'A Proposed model for data warehouse ETL processes', Journal of King Saud University - Computer and Infromation Sciences, Volume 23, Issue 2, July 2011, pp. 91-104

Foshay, N., Mukherjee, A. & Taylor, A. 2007, 'Does data warehouse end-user metadata add value?', *Communications of the ACM*, Volume 50, No.11, November 2007, pp. 70-77

Gangadharan, G. R., & Swami, S. N. 2004,'Business intelligence systems: design and implementation strategies', In *Information Technology Interfaces, June 2004. 26th International Conference*, pp. 139-144, IEE

Gardner, S.R. 1998, 'Building the Data Warehouse', *Communications of the ACM,* Volume 41. No.9, September 1998

Gendron, M. 2014, Business intelligence and the cloud : strategic implementation guide, John Wiley & sons, ISBN: 978-1-118-85984-1 (epub)

Gessner, G. H., & Volonino, L. 2005, 'Quick response improves returns on business intelligence investments', *Information systems management*, *22*(3), 66-74.

Gurjar, Y. S., & Rathore, V. S. 2013, 'Cloud business intelligence–is what business need today', *International Journal of Recent Technology and Engineering*, *1*(6), 81-86.

Hocevar, B. & Jaklic, J. 2010, 'Assessing benefits of business intelligence system - A case study', *Journal of Contemporary Management Issues,* Volume 15, Issue 1, pp.87-119

Inmon, W.H. 2002, Building the Data Warehouse third edition, Wiley Computer publishing, USA, ISBN: 0-471-08130-2

Inmon, W.H. 1996, Building the Data Warehouse second edition, Wiley Publishing, USA, New York

Ives, B., Olson, M. H., & Baroudi, J. J. 1983, 'The measurement of user information satisfaction', *Communications of the ACM*, *26*(10), 785-793.

Jashapara, A. 2004, Knowledge Management: An Integrated Approach, Essex: Prentice Hall

Kelly, M. 1993, 'Assessing the Value of Competitive Intelligence', *Journal of AGSI*, November 1993

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J. & Becker, B. 2008, The Data Warehouse Lifecycle Toolkit Second Edition, Indianapolis, John Wiley & Sons

Khan, R. A. & Quadri, S. M. K. 2014, 'Business intelligence: An Integrated Approach', International Journal of Management and Innovation 6.2.2014, pp. 21-31

KPMG report: Fisher, B., Rast, C., Toplansky, M & Miller, A. 2015, Data and Analytics: A New Driver of Performance and Valuation, June 2015

Krawatzeck, R, 2015, 'Agile Business Intelligence: Collection and Classification of Agile Business Intelligence Actions by Means of a Catalog and a Selection Guide', Information systems management, Volume:32, Issue:3, pp. 177-191

Laskey, K. B., & Laskey, K. 2009, 'Service oriented architecture', *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), 101-105.

Luhn, H. P. 1958, Business intelligence systems. *IBM Systems J.*, 2(3):314–319

Loshin, D 2013, Business Intelligence: The Savvy Manager's Guide, Elsevier 2013, ISBN: 978-0-12-385889-4

Lönnqvist, A. & Pirttimäki, V. 2006, 'The measurement of business intelligence', *Information Systems Management*, Volume 23(1), 32

March, S.T. & Hevner, A.R. 2005, *Integrated decision support systems: A data warehousing perspective*, Decision support systems, Volume 43, Issue 3, April 2007, Elsevier

Martinsons, M.G 1994, 'A strategic vision for managing business intelligence', *Information Strategy: The Executive Journal*, Vol. 10 Issue 3

Martinsons, M.G 1993, 'Towards a Framework for a Strategic Business Intelligence System', Hong Kong: City Polytechnic,

Morton, J., Runciman, B. & Gordon, K. 2014, Big data: Opportunities and challenges, BCS Learning & Development Limited, ISBN: 9781780172620 (eBook)

Moss, L.T. & Atre, S. 2003, Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications, Addison-Wesley Professional, ISBN: 0201784203

Ng, R., Arocena, P., Barbosa, D., Carenini, G., Gomes, L., Jou, S., Leung, R., Milios, E., Miller, R., Mylopoulos, J., Pottinger, R., Tompa, F. & Yu, E. 2013, Perspectives on Business Intelligence, Morgan & Claypool, ISBN: 9781627050944

Olszak, C.M. & Ziemba, E. 2012, 'Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of upper Silesia, Poland', *Interdisciplinary Journal of Information, Knowledge, and Management*,Volume 7, pp. 129-150

Olszak, C.M & Ziemba, E. 2007, ‘Approach to Building and Implementing Business Intelligence Systems’, *Interdisciplinary Journal of Information, Knowledge and Management,* vol. 2, p. 135-148

Przybylek, A 2014, A business-oriented approach to requirements elicitation, 9th International Conference on Evaluation of Novel Approaches to Software Engineering, IEEE conference publications, ISBN: 9897580654

Qian, L., Luo, Z., Du, Y., & Guo, L. 2009, 'Cloud computing: an overview', In *Cloud Computing* pp. 626-631, Springer Berlin Heidelberg.

Reshi, Y.S. & Khan, R.A. 2014, 'Creating Business Intelligence through Machine Learning: An Effective Buisness Decision Making Tool', *Information and Knowledge Management*, Volume 4, No. 1

Saarinen, T. 1996, 'An expanded instrument for evaluating information system success', *Information & Management*, Volume 31(2), pp. 103-118,

Schmarzo, B. 2013, Big data: Understanding how big data powers big business, John Wiley & Sons, ISBN: 9781118740033 (eBook)

Tapadinhas, J. 2014, How to Architect the BI and Analytics Platform, Gartner inc. 2014

Turner, M., Budgen, D., & Brereton, P. 2003, 'Turning software into a service', *Computer*, *36*(10), 38-44.

Yeoh, W 2011 August, ‘Business intelligence systems implementation: testing a critical success factors framework in multiple cases’, *International Journal of Business Information Systems,* Vol.8 (2), p.192

Yeoh, W. & Koronios, A. 2010, 'Critical success factors for business intelligence systems', *Journal of Computer Information Systems,* Volume 50, Issue 3, pp. 23-32

Yi, S.J., Kang, Y.A., Stasko, T.J, *Member, IEEE &* Jacko, J.A. 2007, 'Toward a Deeper Understanding of the Role of Interaction in Information Visualization', IEEE transactions on visualization and computer graphics, Vol. 13, No. 6, November/December 2007