TAMPERE UNIVERSITY OF TECHNOLOGY

MIKAEL RUOHONEN

INCREASING SALES BY A COMPREHENSIVE USE OF CUSTOMER INFORMATION AND ADVANCED ANALYTICAL APPLICATIONS

Master of Science Thesis

Prof. Mika Hannula has been appointed as the examiner at the Council Meeting of the Faculty of Business and Technology Management on December 8, 2010.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Information and Knowledge Management

RUOHONEN, MIKAEL: Increasing sales by a comprehensive use of customer information and advanced analytical applications
Master of Science Thesis, 108 pages, 3 appendices (3 pages)
April 2011
Major: Business Information Management
Examiners: Professor Mika Hannula and Master of Science Riku Lindfors
Keywords: Customer analytics, business analytics, business intelligence architecture, customer insight, customer data, analytical applications and tools.

Companies are required to understand their customers more in depth in order to answer to the challenges introduced by the growingly complex operating environment. This understanding can be acquired through customer analytics in which the available customer information is analyzed with the help of advanced analytical applications. This research studied both customer analytics and the business intelligence architecture required to make customer analytics work. The aim of this study was especially to identify the correlation between the business intelligence architecture maturity and the insightfulness of customer analytics. In addition, particularly the application areas of customer analytics producing customer insight, which can be used to increase sales or sustain current sales, were focused on.

The research was conducted as a case study including five different case companies. A semi-structured interview was used as a data collection method. Additionally, case descriptions including both the current status of business intelligence architecture and customer analytics in the case companies were created based on these semi-structured interviews. Furthermore, the case descriptions were analyzed in order to evaluate the business intelligence architecture maturity, amount of different application areas of customer analytics, and the level of customer analytics' sophistication in the case companies. The results of these analyses were then compared to each other creating understanding from the correlation between these three entities.

Based on these results a conclusion was drawn that there exists a correlation especially between the use of comprehensive customer information and advanced analytical applications and the insightfulness of company's customer analytics. Furthermore, there also exists a correlation between the insightfulness of the company's customer analytics and its ability to use customer information to further increase sales. The main results of this study can be used as a guideline when developing business intelligence architecture and as a source of ideas for new application areas of customer analytics.

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO
Tietojohtamisen koulutusohjelma

RUOHONEN, MIKAEL: Myynnin kasvattaminen kokonaisvaltaisen asiakastiedon ja kehittyneiden analyyttisten sovellusten käytön avulla
Diplomityö, 108 sivua, 3 liitettä (3 sivua)
Huhtikuu 2011
Pääaine: Tiedonhallinta
Tarkastajat: Professori Mika Hannula ja kauppatieteiden maisteri Riku Lindfors
Avainsanat: Asiakasanalytiikka, liiketoiminta-analytiikka, liiketoimintatiedon hallinnan arkkitehtuuri, asiakasymmärrys, asiakastieto, analyyttiset sovellukset ja työkalut.

Yritysten täytyy ymmärtää asiakkaitaan entistä syvällisemmin kyetäkseen vastaamaan yhä monimutkaisemman liiketoimintaympäristön aikaansaamiin haasteisiin. Tällainen asiakkaiden syvällisempi ymmärtäminen on mahdollista saavuttaa asiakasanalytiikan avulla. Asiakasanalytiikassa asiakastietoa analysoidaan kehittyneiden analyyttisten sovellusten avulla. Tämä tutkimus tutki sekä asiakasanalytiikkaa että liiketoimintatiedon hallinnan arkkitehtuuria, jonka asiakasanalytiikka vaatii toimiakseen. Tämän tutkimuksen tavoite oli tunnistaa liiketoimintatiedon hallinnan arkkitehtuurin maturiteetin ja asiakasanalytiikan oivaltavuuden välinen riippuvuussuhde. Lisäksi tutkimuksessa keskityttiin etenkin niihin asiakasanalytiikan sovellusalueisiin, jotka tuottavat asiakkaisiin liittyvää ymmärrystä, mitä voidaan käyttää myynnin kasvattamiseen tai nykyisen myynnin ylläpitämiseen.

Tämä tutkimus suoritettiin viiden yrityksen tapaustutkimuksena. Tiedonkeruumenetelmänä tutkimuksessa käytettiin teemahaastattelua. Tämän lisäksi jokaisen teemahaastattelun pohjalta kirjoitettiin tapauskuvaukset, jotka käsittelivät yritysten liiketoimintatiedon hallinnan arkkitehtuurin ja asiakasanalytiikan nykytilaa. Sen lisäksi näitä tapauskuvauksia analysoitiin, jotta yrityksien liiketoimintatiedon hallinnan arkkitehtuurin maturiteettia, erilaisten asiakasanalytiikan sovellusalueiden määrää sekä asiakasanalytiikan hienoistuneisuuden tasoa kyettiin arvioimaan. Tehtyjen analyysien tuloksia verrattiin toisiinsa, jotta näiden kolmen eri kokonaisuuden välistä yhteyttä kyettiin ymmärtämään.

Näistä tutkimuksen tuloksista voitiin päätellä, että erityisesti kokonaisvaltaisen asiakastiedon ja kehittyneiden analyyttisten sovellusten käytön sekä asiakasanalytiikan oivaltavuuden välillä on riippuvuussuhde. Tämän lisäksi myös yrityksen asiakasanalytiikan oivaltavuus ja yrityksen kyky käyttää asiakastietoa myynnin kasvattamiseen liittyvät toisiinsa. Tämän tutkimuksen päätuloksia voidaan käyttää sekä ohjenuorana kehitettäessä liiketoimintatiedon hallinnan arkkitehtuuria että ideoiden lähteenä uusia asiakasanalytiikan sovellusalueita mietittäessä.

PREFACE

This Master of Science Thesis was done based on a subject received from International Business Machines Oy Ab (IBM). Additionally, also the funding for this study was received from IBM. The study was supervised and examined by Professor Mika Hannula and Riku Lindfors who works as a business manager in Business Analytics and Optimization service line in IBM. I would like to thank both of these gentlemen from the comments and advices they gave me during this project. These comments and advices helped me to focus this study to a correct direction. Additional thanks belong to my colleagues who helped me especially in contacting and finding good case companies for the study. And naturally, I want to also thank the case companies which participated to this study.

Especially I want to thank Sonja Lavonen. She has supported me a lot during this long project and additionally she helped me and the readers of this study by proofreading the study report. Thank you very much for that.

Helsinki, 20.04.2011

Mikael Ruohonen

TABLE OF CONTENTS

ABBREVIATIONS AND TERMS

BA	Business Analytics includes the extensive use of data, different types of analysis and models for providing insight which guides business decisions (Davenport & Harris 2007, p. 7).
BI	Business Intelligence includes architectures, applications, databases, methodologies and tools needed to transform data into information that can be used to make smarter decisions which result to better actions (Turban et al. 2008, p. 9).
Business intelligence architecture	Business intelligence architecture includes data sourcing, data integration and quality, data warehouse architecture and scope, analytical technologies, and business intelligence technologies.
CRM	Customer Relationship Management is a system that includes all the aspects of company's interactions with its customers (Cunningham 2002, p. 6, 112).
Customer analytics	Customer analytics includes the application of advanced analytical methods to the customer data of a company in order to help business in identifying, attracting, and retaining the best and most profitable customers (Schroeck 2001, p. 4).
Customer insight	Customer insight is a deep understanding of customer behavior and needs which can be used to make decisions.
ETL process	Extract, Transform, and Load process is a form of data integration which is often used when data is integrated from data sources into data warehouse.

KPIs	Key Performance Indicators are predefined indicators of business performance which can be used for example in reporting.
OLAP tools	Online Analytical Processing tools are tools which can be used to analyze multidimensional data.
TDWI	The Data Warehousing Institute is and educational institute of business intelligence and data warehousing for executives and professionals worldwide.

1. INTRODUCTION

The customer segments of today are a heterogeneous mass for many industries and thus companies face challenges when defining proper actions needed for certain customers. This is a result of changing demographics and value systems which have fragmented customer value drivers. Also the customer touchpoints such as email, portals, user communities, etc. have multiplied and therefore companies now have much more information about their customers. Because of the many customer touchpoints and different systems handling customer related information the customer data can also be very scattered throughout the organization. Both the amount and disjointedness of customer data add complexity in to the processes of using customer related data.

In all industries, organizations have challenges with delivering a consistent, non repetitive and personalized customer experience across the enterprise and all touchpoints. They also have challenges with acquiring a complete customer picture including all the relevant transactions, interactions and characteristics related to the customer. On top of that it is hard to understand customer behavior and truly be proactive in customer related activities.

In order to answer to this growingly complex operating environment companies require a better understanding about their customers. They need to able to form a complete picture containing all the relevant information about their customers and analyze that information. Through customer analytics and customer insight companies will be able to respond easier to the critical needs such as need for customer loyalty, revenue assurance, and growth. It is highly beneficial to analyze the customer information in order to for example find patterns which can be used to identify cross-selling and up-selling opportunities, “moments of truth” when customer related activities have high possibility to have an impact, customers which are about to go to the competitors, and so on.

1.1. Research objectives and limitations

The focus of this research is to study the different ways customer analytics can be applied in a company to provide new insight and identify how this insight can be used to improve the customer related activities such as sales, marketing, and so on so that the company’s sales is increased. This is studied especially from the perspective of customer information and advanced analytical applications.

Therefore the main objective of this research is to prove the following hypothesis:

The use of comprehensive customer information and advanced analytical applications creates insight which can be used to further increase sales.

Thus the main research question is:

Does the use of comprehensive customer information and advanced analytical applications create insight which can be used to further increase sales?

This research question has been further divided into five sub-questions which need to be answered in order to achieve the research objective and to answer the main research question. These sub-questions are:

1. What is business analytics and how is it used in business?

2. What kind of business intelligence architecture business analytics requires?

3. How can business analytics be applied to customer information and what kind of insight can be obtained with customer analytics?

4. How can customer insight be utilized to increase company's sales?

5. How does the maturity of different aspects of business intelligence architecture affect the insightfulness of customer analytics?

The first four sub-questions will be answered with a review of earlier researches as they are theoretical in nature. In addition, the third and four questions are also discussed in the empirical part as the research material includes some additional answers for these questions. The last, fifth, question will be answered with the help of the findings obtained in the theoretical part and the material collected for the empirical part.

Analytical capability includes multiple dimensions which all need to be kept in mind when business analytics is studied. Davenport et al. (2010, p. 186-188) state that these dimensions are accessible and high-quality data, enterprise orientation, analytical leadership, strategic targets, and analysts. Even though company's analytical capability consists of these many factors this research focuses on to study analytical capability from the point of view of information technology, or more precisely business intelligence, architecture. Therefore especially the dimensions related to enterprise orientation, analytical leadership and analysts are scoped out of the research and the focus is on technology-oriented subjects. Additionally, this means that the focus of this study is to show the correlation between business intelligence architecture and the ability to increase sales. When only part of the multiple different dimensions related to analytical capability are focused on, one cannot unambiguously prove the causal relationship between the two.

In addition the mentioned strategic targets dimension is only partly covered in this study. Firstly, only the different application areas of business analytics based on customer information are studied. This limitation scopes out such application areas as supply chain analytics, spend analytics, and so on. Secondly, especially the application areas of customer analytics which sustain current sales or increase sales directly or indirectly are studied. This leaves out application areas such as fraud and risk analytics.

Furthermore, the focus of this research is on business analytics and advanced analytical applications and therefore the traditional business intelligence or information access and reporting is mainly scoped out. However, in order to define business analytics also business intelligence and the differences between these two concepts must be defined.

1.2. Research approach

There exists multiple different research approaches in literature. A commonly used classification in business economics is defined by Neilimo & Näsi (1980, p. 31). They have classified four different research approaches by dividing the use of information between descriptive and normative and by dividing the methods of information gathering between theoretical and empirical. These research approaches are conceptual, nomothetical, decision-oriented, and action-oriented. On top of these four research approaches Kasanen et al. (1993, p. 255) have defined a fifth approach, constructive approach. The combined classification including all the five approaches is illustrated in Figure 1.

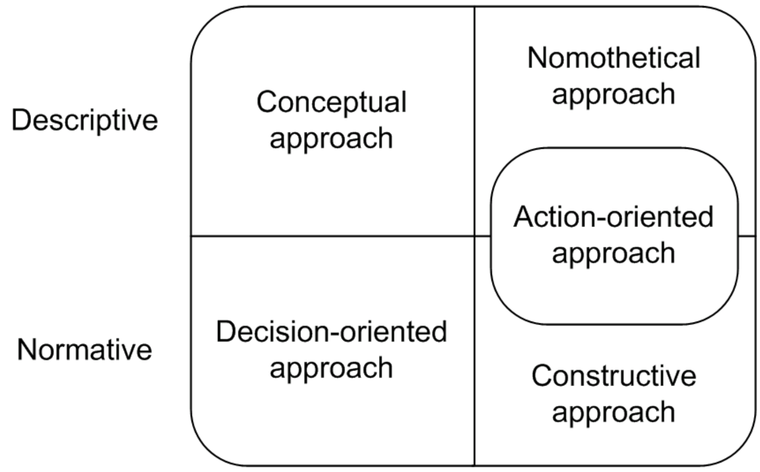

Figure 1. Research approaches used in business economics (adapted from Kasanen et al. 1993, p. 257).

This study can be clearly divided into theoretical and empirical parts. The first part is theoretical where the key concepts of the study are explored by a literature review. The theoretical part follows the conceptual approach. Olkkonen (1994, p. 65) defines that conceptual approach develops conceptual systems, which can be used to describe, identify, or categorize different kind of phenomenon. The purpose of these conceptual systems is to serve some specific purpose, task, or other defined need. Therefore the objective of theoretical part is to get a preliminary understanding of the research area and to build a framework for conducting the empirical research on the case companies.

The empirical part of this study uses the action-oriented approach. In action-oriented approach the objective is to understand a phenomenon and possibly to develop a new theory (Olkkonen 1994, p. 76). In action-oriented approach empirical analysis is often quite profound study of few selected companies (Näsi 1980, s. 31). And as this research will focus on a few selected companies in order to understand a phenomenon the research method is a multiple case study. Case study is an empirical inquiry which's purpose is to examine a contemporary phenomenon within its real-life context. It should be used especially when the boundaries between phenomenon and context are unclear. Multiple different sources of evidence and data-collection methods can be used in a case study. These sources of evidence include documents, archival records, interviews, direct observation, participant-observation, and physical artifacts. (Yin 2003, p. 13.) The selected research method is further described in chapter 5.

1.3. The structure of the thesis

The research can be divided into theoretical and empirical parts. This division is illustrated in Figure 2. The theoretical part includes three chapters which are all based on to the previous research from this particular field of study. The used source material includes large amount of different scientific literature, publications, articles, journals, and additionally some white papers which supported the other source material.

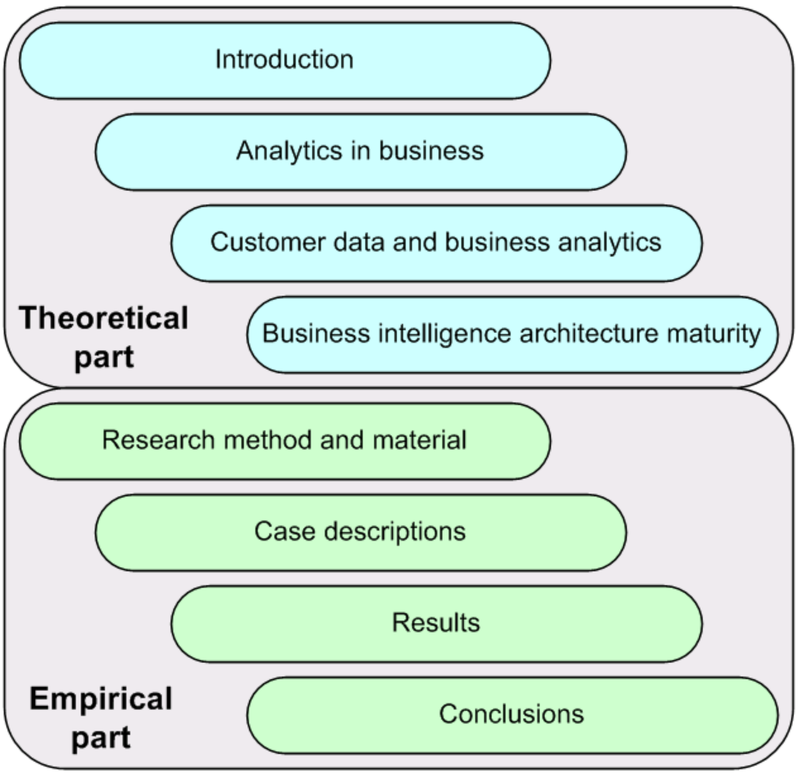

Figure 2. The structure of the thesis.

The theoretical part starts with a chapter (chapter 2) which describes what analytics means in the context of business. Such main concepts as information, information management, business intelligence, business analytics, and business intelligence architecture are defined in this chapter. The following chapter (chapter 3) on the contrarily describes the concept of customer analytics in detail. This includes also the definition of customer data and related concepts of customer analytics. In addition, the last chapter of theoretical part (chapter 4) describes different business intelligence and business analytics related maturity models and composes a new business intelligence architecture related maturity model which will be used to evaluate the case companies in the empirical part.

Furthermore, the empirical part includes four chapters which all target ultimately to answer the research questions and to prove the research hypothesis. The first chapter of the empirical part (chapter 5) defines the research method and material that is used in this research. The following chapter (chapter 6) on the other hand describes all the case companies based on the semi-structured interviews which were held to the representatives of each case company. In the next chapter (chapter 7) the case descriptions are analyzed and the results of the research are presented. Also the assessment of the empirical part of the study is included into this chapter. And lastly, the main results of the research as well as the implications to practice and theory are described in the last chapter of the empirical part (chapter 8). This chapter includes also the assessment of the whole study and proposals for further research.

2. ANALYTICS IN BUSINESS

In this chapter the role of analytics in business is discussed. In order to do that information and information management are described in the first section. Then, in the second section, the concepts of business intelligence and business analytics are introduced. In addition, also the differences between these closely related concepts are defined. And lastly the business intelligence architecture which is required for business analytics is described in the third section.

2.1. Information and information management

This section describes the different levels and dimensions of information in order to create an understanding from the concept of information. Additionally also the process of information management is introduced so that a basic idea from the different steps related to the use of information can be acquired.

2.1.1. *Levels of information*

In order to understand business intelligence and business analytics it is important to understand the relationship between *data*, *information*, *knowledge* and *intelligence*. This relationship has been illustrated in Figure 3.

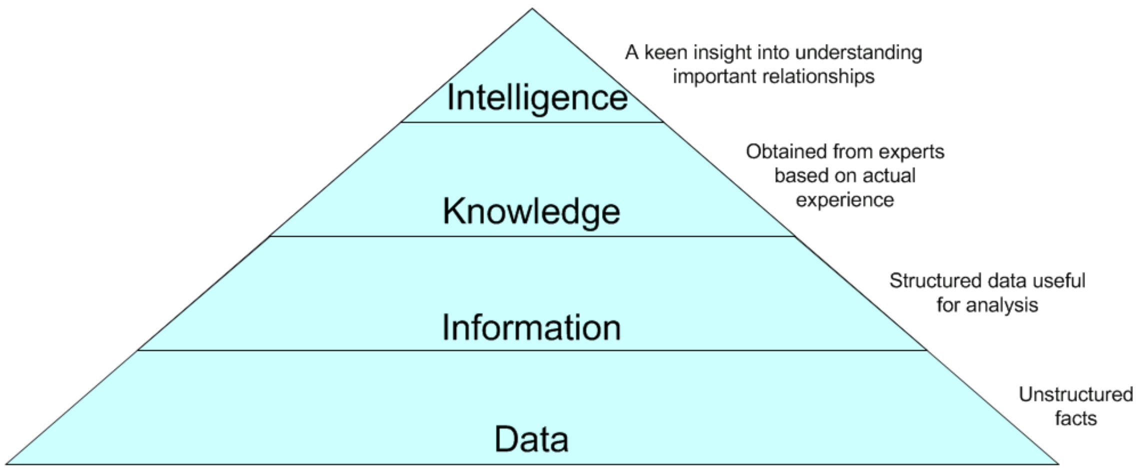

Figure 3. Levels of information (adapted from Thierauf 2001, p. 8).

Thierauf (2001, p. 7) states that data in this context represents unstructured facts and figures which are not very helpful in their current form for the decision makers. Additionally, Turban et al. (2005, p. 214) argue that data is recorded, classified, and stored items about activities, events, things, and transactions. The main characteristic of such data is that it is not organized to convey any specific meaning. For example a single record in an operational database that records the 21.35€ purchase of a Shakespeare book from a website by a customer from Helsinki, Finland is an example of data.

Information, described as the next level in the illustration, is structured data and therefore it is more useful for decision making and analysis (Thierauf 2001, p. 8). When data is endowed with a degree of business context and organized in such a manner that it is meaningful for a recipient it can be called information. This information can either reveal something not known for the recipient or it can confirm something that is already known. (Turban et al. 2005, p. 214-215.) Information can for example be data that has been filtered, sorted, synthesized, and aggregated (Liautaud & Hammond 2001, p. 5). A list of Shakespeare purchases by customers with similar demographic attributes is an example of information.

Information is transformed in to knowledge when a range of information is integrated by an expert in order to find trends and patterns that enable further insight and prediction (Thierauf 2001, p. 9). Turban et al. (2005, 215) state that in order to form knowledge data items and information need to be organized and processed in such a way that they convey experience, expertise, understanding, and accumulated learning. Furthermore, Thierauf (2001, p. 9) argues that information is data about the data and that knowledge is information about the information.

Consequently, intelligence represents the comprehensive understanding about larger context, for example about company's customers and operations (Thierauf 2001, p. 9). Additionally, insight is also a widely used term in literature related to business intelligence and especially business analytics. Thierauf (2001, p. 8) defines that intelligence is a keen insight into understanding important relationships. Davenport et al. (2010, p. 6-7) argues that for example insight into the past explains how and why something happened, insight into the present provides recommendations for current actions, and insight into future predicts best possible future actions in order to reach best possible results. Insight is more widely discussed in section 2.2.1 where the differences of business intelligence and business analytics are described.

Overall when thinking the difference between different levels of information one can notice that the focus of data is typically operational. Example of such operational data could be the sales calls of salespersons. Information in the same context would be for example the monthly summary reports from these sales calls. Thus information has a more tactical focus. In turn, the understanding of theses summary reports and their changes over a longer period of time can be considered knowledge. However, when the future direction and emphasis of company's sales activities is set there is need for insightful analyses based on the different levels of information acquired from the sites and functions of the company. As a result of such analyses intelligence is created.

Therefore knowledge and especially intelligence has a strategic focus. (Thierauf 2001, p. 10.)

2.1.2. *Dimensions of information*

After defining the relationship between different levels of information one needs to consider the dimensions of business information in order to understand information needs of an organization. Hannula & Pirttimäki (2005, p. 37) argue that business information has three dimensions which are needed for definition of information needs and enriching information. These dimensions are *information subject*, *information source* and *information type*. Figure 4 illustrates these dimensions and the two different aspects each dimension has.

In this figure the information subject dimension illustrates the subject that the information is about. On the different sides of the dimension there are organization's internal and external subjects. (Hannula & Pirttimäki 2005, p. 38.) Internal subjects can be for example resources, work force or capital. External subjects on the other hand can be for example customers and competitors. Secondly, there are multiple information sources where an organization can acquire information. Information source dimension represents the two high level categories for information sources, internal and external (Hannula & Pirttimäki 2005, p. 38). Organization's operative systems are good examples of internal sources, whereas Internet is an example of external source. The third dimension, information type, defines whether the information is quantitative or qualitative. Quantitative information is easy to capture, measure, process and communicate. Qualitative information is instead hard to measure, compare and communicate forward as it is descriptive. (Hannula & Pirttimäki 2005, p. 39.)

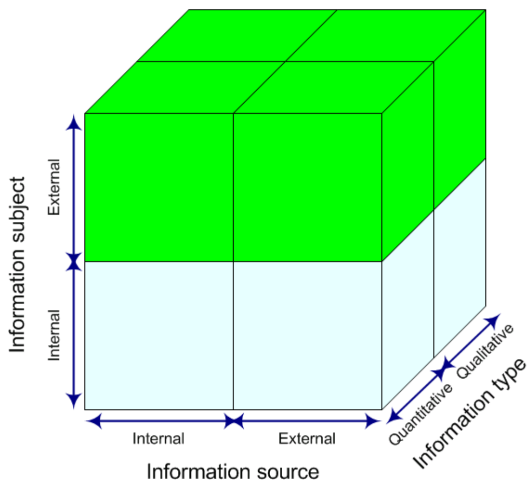

Figure 4. Relevant dimensions of business information for this research (partly adapted from Hannula & Pirttimäki 2005, p. 38).

In other words, Hannula & Pirttimäki (2005, p. 38) define that the cube of business information can be used as an illustrative tool to define different organizational information needs. In this research the cube of business information is used to understand what dimensions are relevant for the research. This is done by highlighting the relevant sub-cubes in Figure 4. This research focuses on customer information which means that the information subject is external. However, as customer information can be acquired both from internal and external sources this research includes both sides of information source dimension. Internal sources include for example transactional systems while external sources include for instance social media. Additionally also both quantitative and qualitative customer information are included in this research. Both internal and external sources can have quantitative and qualitative customer information. Although it must be stated that more often for example different systems or data warehouses include mainly quantitative data. However, organizations store also documents and other unstructured content in their systems and therefore depending on the case in hand information type can be also qualitative. A good example from a technique that uses qualitative data is for instance text mining. Text mining and other techniques will be discussed in more detail in section 2.3.4.

2.1.3. *Process model of information management*

On top of understanding the differences between the levels of information it is also beneficial to understand in high level how information can be managed. The process model of information management addresses the steps required to the transformation of data into information which can be then used as a basis for better decisions. In his widely accepted model Choo (2002, p. 24) defines information management as a continuous cycle consisting of six activities. These closely related activities are identification of information needs, information acquisition, information organization and storage, development of information products and services, information distribution, and information use. Information use is finally followed by the creation of new information which also starts the cycle. New information is created by organizations adaptive behaviour. (Choo 2002, p. 24.)

Other similar models are defined widely in scientific literature. For example Pirttimäki (2007, p. 74) introduces a generic business intelligence process including five phase. These phases are specification of information needs, gathering of information, processing of information, dissemination of information, and utilization of information. Additionally, for instance, Davenport & Prusak (1997, p. 135) define a generic information management process which consists from four steps. These steps are determining information requirements, capturing information, distributing information, and using information. The actual model considered is not however that important. The thing which on the other hand is important especially from the perspective of business intelligence and business analytics is the fact that information needs to be gathered, processed, distributed, and used. It is not enough for example to just gather information. Information needs to be further processed, distributed to where it matters, and acted upon.

2.2. Business intelligence and analytics

The goal of this section is to introduce the concepts of business intelligence and business analytics. These two concepts are discussed in detail in order to create an understanding from the differences between these two concepts. Additionally also the different types and applications areas of business analytics are described shortly.

2.2.1. Difference between business intelligence and business analytics

Business Intelligence (BI) has been conceptually around all the way from the Management Information Systems (MIS) in 1970s (Turban et al. 2008, p. 9). The basic idea behind business intelligence is to transform data into information that can be used to make smarter decisions which result to better actions. The term includes architectures, applications, databases, methodologies and tools needed to reach this transformation. (Eckerson 2007b, p. 5; Turban et al. 2008, p. 9.)

Gartner (2009, p. 14) defines that business intelligence encompasses the people, processes, applications, and tools which are used to organize information, enable access to information, and to analyze information for better decision making and performance management. Therefore it is important to remember that also people are an important part of business intelligence even though that specific aspect is scoped out of this research. Additionally, Gartner (2009, p. 14) defines also that business intelligence focuses on locating and accessing the information needed to the company's analytical, business and decision processes. Business intelligence presents, through efficient processes and technologies, the needed information in the most usable formats (Gartner 2009, p. 14).

Analytics on the other hand began to command attention in business as early as the late 1960s when computers were started to use to analyze data in decision support systems (DSS). These systems were used in a few data-intensive business functions such as production planning and transportation routing for analytical and repetitive tasks. After that analytics have evolved with the development of executive support systems, enterprise resource planning (ERP) systems, data warehouses, and other systems. (Davenport & Harris 2007, p. 11-12.)

Literally analytics is the science of analysis, which generally can be interpreted to refer to analysis of data (Turban et al. 2008, p. 86). Depending on the context, the data and the analysis methods as well as the used tools can be different. Davenport & Harris (2007, p. 7) define that analytics means in the context of business the extensive use of data and different types of analysis and modeling done based on this data. In other words *business analytics* (BA) is the use of data and models for providing insight which guides business decisions. This analysis and modeling can include for example statistical and quantitative analyses as well as explanatory and predictive models.

Furthermore, analytical initiatives usually happen in support of larger business intelligence and performance management efforts (Gartner 2009, p. 14).

As discussed, business intelligence can be defined as the architectures, tools, technologies, methodologies and processes needed to turn data into information and information into knowledge and plans that optimize business actions (Eckerson 2007b, p. 5; Turban et al. 2008, p. 9). Furthermore, business intelligence architecture encompasses among other things data integration, data warehousing, and reporting and analytical tools (Davenport & Harris 2007, p. 156-158).

However, even though both reporting and analytical tools are part of business intelligence architecture there exist significant differences in the questions or problems that reporting tools and analytical tools answer (Turban et al. 2005; Davenport & Harris 2007; Jaspersoft 2010). Multiple sources make distinction between traditional business intelligence functionality and business analytics. Standard reporting, queries, ad hoc reporting, alerts, and so on are seen as traditional business intelligence functionality which covers the reporting and data analysis needs of business users. (Turban et al. 2005; Davenport & Harris 2007; IBM 2009a; Jaspersoft 2010.)

Business analytics on the other hand is considered to include the more advanced domains traditionally associated with analytics such as statistical analytics, data mining, predictive analytics, optimization, simulation, and rule engines (Turban et al. 2005; Davenport & Harris 2007; IBM 2009a; Jaspersoft 2010). Therefore business intelligence and related tools are discussed in the context of accessing data and simple analysis of data. On the contrary, business analytics and related tools are discussed in the context of uncovering patterns within large volumes of data in order to predict behaviour and anticipate future.

In this research business analytics is defined as a continuum of business intelligence. In many cases business analytics comes into the picture when information presented through business intelligence technologies raises questions and a need to analyze the situation further. The differences and connections between business intelligence and business analytics are discussed in more detain in the coming subsections.

Competitive advantage and degree of intelligence

Competitive advantage and degree of intelligence are both aspects related to reporting and analytics. Specific types of reporting or analytics produce a specific degree of intelligence and therefore allow certain competitive advantage. (Davenport & Harris 2007, p. 7.) The difference in competitive advantage and degree of intelligence between business intelligence and analytics has been illustrated in Figure 5.

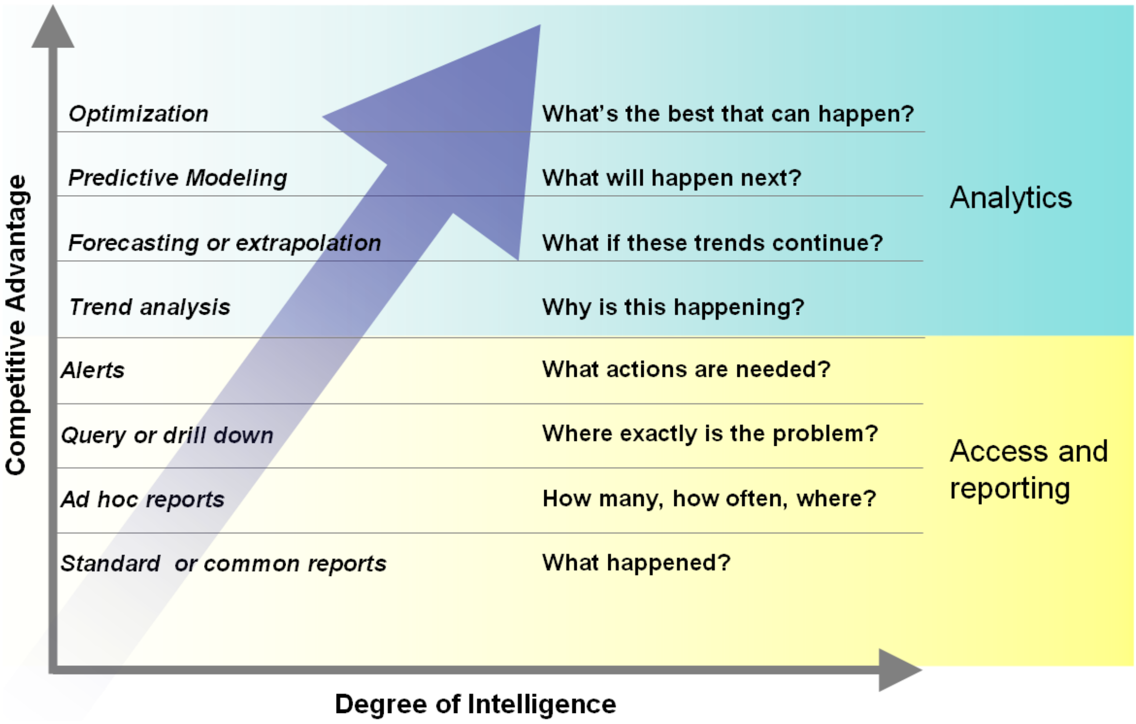

Figure 5. Business intelligence & analytics (adapted from Davenport & Harris 2007, p. 8).

Business intelligence traditionally focuses on data access and reporting, demonstrated by the lower part of the Figure 5. It uses a consistent set of metrics to both measure past performance and guide business planning (Davenport & Harris 2007, p. 7-8). Business intelligence technologies provide functionality such as querying, reporting, dashboards, scorecards, and alerts based on exceptions. All these technologies examine what happened in the past and are deductive in nature. This means that the end users need personal experience about the patterns and relationships that exist within the data. They use business intelligence technologies for accessing and exploring the data in order to validate their own hypotheses. Some of the business intelligence technologies such as dashboards and scorecards use predefined metrics and key performance indicators (KPIs) which take the deductive reasoning a little further. End users only have to monitor these metrics and KPIs on a regular basis. (Eckerson 2007b, p. 6.) Business intelligence gives answers to the following kind of questions: What happened? How many, how often, where? Where exactly is the problem? What actions are needed? (Davenport & Harris 2007, p. 8.)

Assuming one would have all the necessary business intelligence technologies available, standard report from sales made in specific region in last quarter can tell us for example what happened in that specific region in last quarter concerning sales. Depending on the content of the report one can start making correcting actions if necessary. As an example, if the standard sales report shows that there has been a sudden decrease in sales, one can for instance create an ad hoc report to find out how many over 1,000€ deals were lost. Additionally one can drill down in to a specific lost deal for finding out the needed facts in order to understand where exactly the problem was. Furthermore, one might now want to make sure that always when a deal over 1,000€ is lost an alert is generated immediately. Such alerts based on predefined conditions can be created by business intelligence technologies.

Analytics on the other hand focuses on using data and statistical methods for developing new insights and understanding of business performance, demonstrated by the upper part of the Figure 5. In order to drive decision making business analytics uses extensively data, statistical and quantitative analysis, extrapolation, explanatory and predictive modeling, data mining, optimization and simulation (Davenport & Harris 2007, p. 169-170). Compared to business intelligence technologies for example predictive modeling is inductive. It employs different predictive methods which explore all the data available to find out meaningful relationships and patterns. Thus it lets data lead the way and does not presume anything about the data. (Eckerson 2007b, p. 6.) Business analytics gives answers to the following kind of questions: Why is this happening? What if these trends continue? What will happen next? What is the best that can happen? (Davenport & Harris 2007, p. 8.) Business analytics may be used as part of automated decision processes or as an input for human decisions.

One can use the above discussed business analytics technologies such as a trend analysis for example for extracting underlying patterns from the sales related data of certain region. This analysis can be used for instance to identify why the amount of the lost deals over 1,000€ has increased in the last two quarters so much. Based on the result of the trend analysis one can now make corrective actions to the right things without having to rely on gut feeling. One can also use forecasting or extrapolation based on the sales related data in that certain region to find out for example what kind of an impact these trends that have led to the increase in the amount of lost over 1,000€ deals will have to the business if they are let to continue. This information can for instance be used as a good reason for the necessary changes in that region's sales. Furthermore, by creating predictive models which predict from the sales related data the deals that have a high possibility to be lost one can focus needed efforts on those deals in order to win as many as possible from them. In addition, with optimization technologies one can for example define based on the sales related data what are the best results which can be achieved with the current sales force and consequently allocate the sales force in to the sales related activities in an optimal way.

Innovation and time frame

The difference between business intelligence and business analytics can also be observed by organizing the fundamental questions organization needs to answer about its business in two dimensions; time frame and innovation (Davenport et al. 2010, p. 6). Figure 6 identifies six important questions that business intelligence and business analytics technologies can address. The questions have been categorized based on weather one is looking at the past, present or future and based on is one working with known information or gaining new insight.

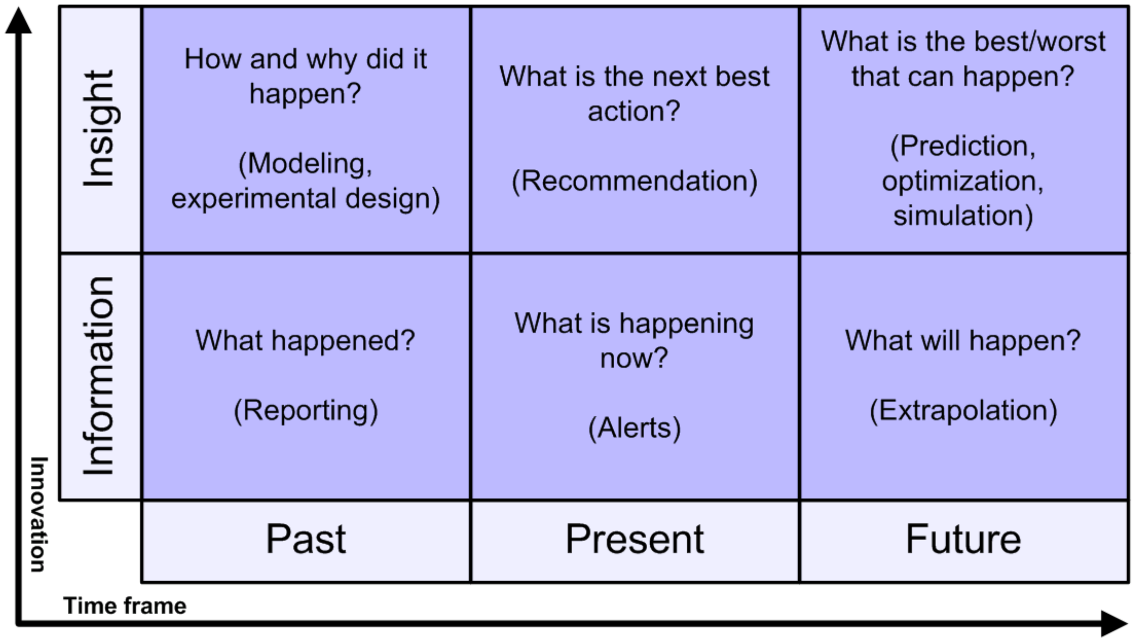

Figure 6. Key questions addressed by business intelligence and business analytics (adapted from Davenport et al. 2010, p. 7).

Even though the difference is flickering one can notice that many of the traditional business intelligence focus areas are in the bottom row of Figure 6. They mainly answer the information-oriented questions and try to solve the challenges in using information more effectively. Reporting, for example, is clearly information-oriented and looking at the past. Alerts on the other hand apply information about normal performance in order to generate alerts about the present performance. Finally for example forecasts can be created by using simple extrapolation on past information. (Davenport et al. 2010, p. 6.)

In order to create new insight different tools digging deeper in to the data are needed. For example statistical modeling gives us new insight by telling how and why certain things happened in the past. Technologies that compare for example customer behaviour and give recommendations for additional possibly interesting products for specific customer produce insight in to the present. Prediction, optimization, and simulation technologies allow us to gain insight into the future by telling us how to reach the best possible future results. (Davenport et al. 2010, p. 6-7.) Therefore the main focus of business analytics can be seen in the top row of Figure 6 where the answers to the fundamental questions tell us why something happens and is it likely to recur.

2.2.2. *Different types and application areas of business analytics*

In general, business analytics can be divided into *explanatory analytics* and *predictive analytics* (Davenport & Harris 2007, p. 7; LaValle 2009, p. 7). Additionally, for instance Rygielski et al. (2002, p. 487) mention *forensic analysis* as a one form of analytics. Turban et al. (2005, p. 266) write that for example data mining tools which are described in more detail in section 0 are used especially for hypothesis-driven and discovery-driven analyses. These both can be seen as explanatory analytics as hypothesis-driven data mining is defined as a validation of user made proposition, and discovery-driven data mining is defined as finding of patterns, relationships, and associations among the data. Furthermore, also Rygielski et al. (2002, p. 487) state that in discovery (explanatory analytics) the data is looked through without predetermined hypothesis in order to find hidden patterns.

Predictive analytics on the other hand is related to the prediction of what is going to happen such as the prediction of stock or customer behavior (Davenport & Harris 2007, p. 178). Also Zaman (2005, p. 1) and Rygielski et al. (2002, p. 488) state that predictive analytics is the branch of data mining, which is related to the prediction of future probabilities and trends. Additionally, Turban et al. (2008, p. 103-104) write that predictive analysis tools use sophisticated algorithms to determine recurring likelihood of different situations and probable future outcomes of specific events. Furthermore, predictive analytics can be used to make complex predictions including multiple variables as for example traffic congestion level predictions. Traffic congestion level prediction requires information about traffic flow, incidents, speeds and locations of vehicles, weather forecasts and conditions, special scheduled events, and so on. The main part of predictive analytics is the variable or variables which can be measured in order to predict future behavior. These variables are called *predictors*. (Turban et al. 2008, p. 104.) Predictive models are created by combining multiple predictors. These models can then be used in analysis with an acceptable level of accuracy to forecast future probabilities. (Zaman 2005, p 1-2.)

Furthermore, as stated, Rygielski et al. (2002, p. 487) add a third type of business analytics; forensic analysis. In forensic analysis the patterns which are extracted from the data are used to find anomalous or unusual data elements (Rygielski et al. 2002, p. 488). Forensic analysis can be used for instance by banks and insurance companies to identify frauds. For example anomalous bank transactions can be identified, investigated, and possibly acted upon with the help of such analyses.

On top of the different types of business analytics there are also multiple different application areas of business analytics (Davenport & Harris 2007, p. 57-58). These include such internal processes as for example general management, manufacturing, finance and accounting, research and development, and human resource management. Furthermore, analytics can be applied also to external processes such as customer relationship management and supply chain management (Davenport & Harris 2007, p. 83). In addition, also for instance Turban et al. (2011, p. 204-206) mention multiple different application areas for analytics. These application areas include among other things customer relationship management, banking and insurance, retailing and logistics, manufacturing and production, brokerage and securities trading, computer hardware and software, government and defense, travel, health care, law enforcement, and sports. Additionally, for example customers and customer relationship management

– the focus areas of this research – are seen as beneficial application areas of business analytics by many other researchers (Schroeck 2001; Rygielski et al. 2002; Fayyad 2003; Lenzen 2004; Voudouris et al. 2008).

2.3. Architecture

As discussed previously, business analytics can be seen as a continuum of business intelligence. Therefore business analytics, as well as business intelligence, requires proper *business intelligence architecture* in order to yield results. In this section a closer look into the business intelligence architecture from the business analytics point of view is taken.

Turban et al. (2008, p. 11) argue that business intelligence architecture includes four fundamental components. These components are data warehouse, business analytics, business performance management, and user interface. Davenport & Harris (2007, p. 156-158) on the other hand state that business intelligence architecture conceptually breaks in to six elements. These elements are data management, transformation tools and processes, repositories, analytic tools and applications, presentation tools and applications and operational processes. Even though the separation is different the context is very much the same. It is also clear that business intelligence architecture is an umbrella term for enterprise-wide set of applications, databases, methodologies, governance processes, and tools, which cover all the necessary things to enable interactive access to data, to enable manipulation of that data, and to give decision makers and business analysts the ability to conduct appropriate analysis (Thierauf 2001, p. 92; Davenport & Harris 2007, p. 155; Turban et al. 2008, p. 9).

The first of the fundamental parts of business intelligence architecture is the data warehousing including the extraction, transformation and loading of data from separate operational systems and data sources, inside and outside the company, in to the data warehouse (Davenport & Harris 2007, p. 165). This is done with transformation tools and processes. Data warehousing also includes the actual data warehouse, a specially designed repository for the transformed and integrated data (Turban et al. 2008, p. 11-12). In the data warehouse the data is organized and enriched with metadata, which is information about the data, in such a way that it supports reporting and analyses (Davenport & Harris 2007, p. 166). Data warehouse could be one large repository or it can for example contain a number of function dedicated data marts which serve the business requirements of a specific company's function or department (Turban et al. 2008, p. 12). The specific details about how a data warehouse is built depend on the company, but the basic principles remain the same in most of the cases. Furthermore, data management has also an important role in data warehousing. Companies require a well defined data management in order to be able to base their reporting and analytics on comprehensive and trusted data (Davenport & Harris 2007, p. 159).

In addition to data warehousing, analytical tools and applications as well as presentation tools and applications are important parts of business intelligence architecture. These tools and applications provide the functionality and methodologies which allow companies to use their data to create meaningful reporting and analytics (Davenport & Harris 2007, p. 167, 171).

Analytical tools and applications of this kind include for example data mining tools which extract patterns from data or rule engines which address logical questions or problems with condition based business rules. These business rules can be combined to other applications in order to create automated processes and recommendations. As another example, some analytical tools and applications are able to use statistical or quantitative algorithms to analyze data so that they find an optimal target; such as an optimal price for a specific product. (Davenport & Harris 2007, p. 168-169.)

Presentation tools and applications on the other hand allow end users to access the analytics provided by analytical tools and applications, and the data stored in data warehouse. These tools and applications provide end users with reporting tools, scorecards, portals, and so on. They also allow end users to create ad hoc reports, share data with their colleges, visualize even complex sets of data, and to receive different kinds of exception alerts. (Davenport & Harris 2007, p. 171.)

To further illustrate the meaning of these concepts it is useful to study other available definitions. Turban et al. (2008, p.12, 83) argue that business analytics in the context of business intelligence architecture means all the software tools that allow the creation of on-demand reports, queries, OLAP (Online analytical processing), data mining, advanced analytical techniques, and so on. After reports and analytics are done based on the data in data warehouses, they need to be delivered and presented to correct people in correct form. Turban et al. (2008, p. 14) state that there are many different visualization tools which perform this task and act as user interface to the reports and analytics. The definition of business analytics provided by Turban et al. can be however quite misleading as there are fundamental differences in accessing and reporting data and analytics as it was defined in section 2.2.1. Furthermore, also Turban et al. (2008, p. 102) discuss advanced business analytics in their book and categorize it primarily as data mining and predictive analytics; thus leaving on-demand reports, queries, and OLAP out of the categorization. This categorization is quite similar as the definition of Davenport & Harris and the definition used in this research. In any case, there exist multiple different analytical and presentation technologies and tools that can be used to perform reporting and analytics depending on the business need in hand. These technologies are further examined in the coming sections.

Furthermore, Turban et al. (2008, p. 14) state that there exists also a business performance management related aspect in business intelligence architecture. Business performance management defines and sets the framework for the measurement of company’s performance. This framework is used when creating reports, queries, and dashboards so that they reflect the certain aspects of performance that company wants to monitor and improve. (Turban et al. 2008, p. 14.) Additionally, things like reliability, scalability, and security are also connected to business intelligence architecture. Operational processes ensure with standards, policies, and processes that organization creates, manages and maintains applications and data in a correct way (Davenport & Harris 2007, p. 172).

Next the main building blocks of business intelligence architecture such as data management, transformation tools and processes, repositories, analytic tools and applications, and presentation tools and applications are described in more detail.

2.3.1. *Data management*

Business analytics is highly dependent on data. Davenport & Harris (2007, p. 158-165) state that one of the most important factors in being prepared for analytics is the sufficient volumes of high quality data. A proper data management strategy ensures that companies which have multiple sources of data such as enterprise resource planning, customer relationship management, point-of-sale systems, and externally gathered data, have the right information and they use it appropriately. Data management strategy needs to consider data sourcing, data relevance, data quantity, data quality, and data governance (Davenport & Harris 2007, p. 158).

Data relevance addresses the value of data. It considers what data is most valuable or needed for analytics so that competitive differentiation and improved business performance can be reached (Davenport & Harris 2007, p. 159). Companies have huge amounts of data and every bit of it cannot be used for analytics. There just are not enough resources in order to do it. It also is not beneficial. Both business and IT needs to work together and define what data is needed for analytics as this definition requires both business and IT understanding (Davenport & Harris 2007, p. 160).

Even though companies can store masses of data it sometimes happens that the needed data for analytics is missing. Either the company never stored that particular data or it never existed. Therefore it is required that companies also try to predict the future data needs. (Turban et al. 2005, p. 218.) Hannula & Pirttimäki (2005, p. 37) write that the needed data can be either about internal or external subject. A relevant external business information subject example for this research is customers, whereas employees are a good example of internal business information subject.

After defining valuable data one needs to know where to find it from. Data sourcing considers where the needed data can be collected. There exist multiple sources for data that companies can use in business analytics. These data sources can be either internal or external (Hannula & Pirttimäki 2005, p. 37). Internal data sources include among other things all the enterprise information systems, transaction systems, and organization’s personal computers and databases with their documents (Davenport & Harris 2007, p. 161). This data includes such things as products, services, processes, sales data, employees, assets, and so on (Turban et al. 2005, p. 215). For example companies that sell information, governments, company’s websites, internet, and market researches are on the other hand external data sources. Also e-mail, CDs, DVDs, RFID tags, and so on can be sources of data. (Turban et al. 2005, p. 215; Davenport & Harris 2007, p. 161.)

Davenport & Harris (2007, p. 162-163) however state that data should not be collected just in case it is needed because it can lead to data overload in which irrelevant data makes finding relevant data harder. Additionally the costs of collecting and storing all available data can out-weight the benefits. Also Turban et al. (2005, p. 220) state that companies need to focus on data which is relevant and value adding. Therefore data management needs to consider how much data is really needed.

In general, companies tend to store their data in functional silos and companies can have multiple separate data sources. Many of the data sources also contain errors, missing values, and integrity problems. Companies also have huge quantities of data. Because of all these things data quality is one of the biggest challenges companies face (Turban et al. 2005, p. 219-220; Davenport & Harris 2007, p. 163). Although challenging, data quality is important as it determines the usefulness of the data and the decisions based on the data (Turban et al. 2005, p. 218). By addressing the different data quality related characteristics companies can increase the value of data and therefore reach better results with business analytics. Data needs to be correct, complete, current, consistent, in context, and controlled (Davenport & Harris 2007, p. 163-164). In other words data needs to be reviewed, monitored, timely for the decision in hand, standardized and common across organization, and it needs to have a defined meaning. Data quality can be addressed both in data sources and during the integration into the data warehouse.

In addition also data governance is part of data management. Data governance addresses those rules and processes which are needed to manage data from its acquisition through its retirement (Davenport & Harris 2007, p. 164). The different stages of data management life cycle are acquisition, cleansing, organization and storage, and maintenance. Data governance needs to determine the following things for successful data lifecycle management. How it is decided what data is needed? How out of date, incorrect, incomplete, or redundant data is detected and removed? How data is systematically extracted, integrated, and synthesized? How and when data is updated? How data security, privacy, and integrity are ensured? How and when data that is not needed is saved, archived, or retired? (Davenport & Harris 2007, p. 164-165.)

2.3.2. *Transformation tools and processes*

In order to get data ready for analytics the data needs to be first extracted from the identified data sources and brought in to a repository (Davenport & Harris 2007, p. 165). Along the way, data from multiple sources needs to be also transformed and cleaned so that the data in the repository is in a correct form. This data integration process is performed with transformation tools and processes.

When data integration is performed in order to support analytical purposes by loading data into the data warehouse for business intelligence and business analytics, ETL (extract, transform, and load) is the most common form of data integration (TDWI 2010, p. 6). This ETL process is illustrated in Figure 7. Of course also other techniques and best practices exist. Data federation, database replication, and data synchronization are techniques used mainly in operational data integration which is used for example between operational systems (TDWI 2010, p. 6).

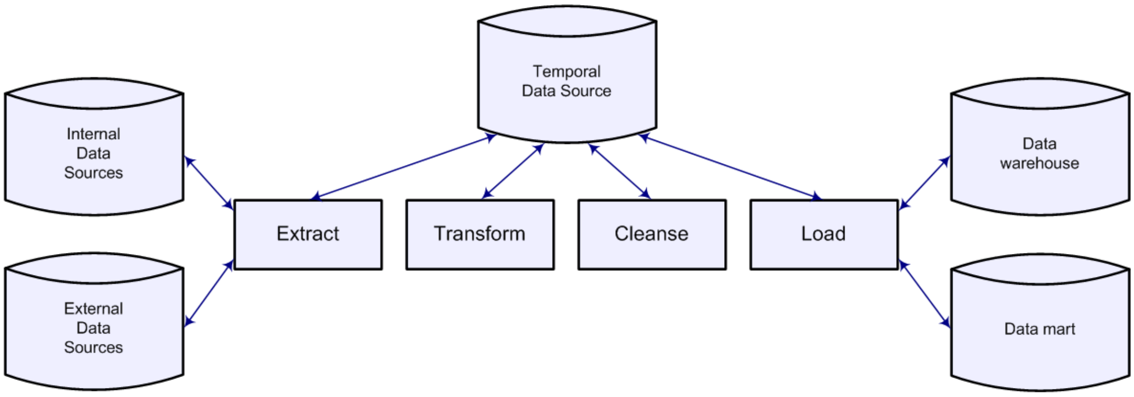

Figure 7. ETL process (partly adapted from Davenport & Harris 2007, p. 165; Turban et al. 2008, p. 55).

The first part of an ETL process involves extracting the data from the source systems so that the transformation of data can be started. In the following transformation and cleansing phases the data needs to be cleansed and validated using business rules (Davenport & Harris 2007, p. 165). Cleansing and validation can include multiple different steps. It is also important to fix any possible data quality issues related to the extracted data if they are not treated in source system level (Turban et al. 2008, p. 55). After all this is done the transformed and cleansed data is loaded into the data warehouse, data mart or both.

The transformation and cleansing phases are important as the extracted data can for example include duplicates especially between data extracted from different sources. These duplicates need to be combined into a single valid record. In transformation phase companies have to also deal with problems related to missing data. In some cases the missing data can be filled with inferred data and in some cases it has to be left blank making it therefore useless for analysis (Davenport & Harris 2007, p. 165-166). Sometimes all the extracted data is not however needed and this obsolete data can be purged. On the other hand all the relevant data needs to be joined and mapped into the new and common format used in the destination system (Davenport & Harris 2007, p. 165). This might require making legitimate field combinations or field divisions as different systems store for example customer address in different level of detail. Also all the existing data harmonization standards need to be applied (Davenport & Harris 2007, p. 165). For example different systems might use different units of measure and therefore there might be a need to harmonize the data so that common units of measure are used. Furthermore, data integrity is one of the major issues related to data quality especially when older systems without all the needed integrity checks are used as source systems. Therefore the integrity issues often need to be covered in the transformation and cleansing phases. In addition, the integrity of existing data in a data warehouse needs to be checked with a series of different methods such as uniformity, version, completeness, conformity, and genealogy checks. (Turban et al. 2005, p. 222.)

There exist specific automated tools which provide the functionality needed to design an ETL process and perform the different steps in ETL processes. However while these tools help in ETL process, considerable manual effort is still needed (Davenport & Harris 2007, p. 165). These tools also typically document how data is changed during the ETL process. This is known as metadata and it is shared with other applications (Turban et al. 2005, p. 224). Also Davenport & Harris (2007, p. 165) state that transformation phase usually includes standardizing business definitions, metadata, related to certain data and concepts. This makes sure that data extracted from multiple sources is comparable.

In addition the timeliness of data is related to data integration. Turban et al. (2011, p. 359) write that traditionally data integration from data sources into data warehouses or data marts has been done for example in weekly or daily basis with overnight batches. They however also argue that often a business cannot wait for instance a whole day in order to have the data available for reporting or analysis. For this reason more real-time data integration is needed. (Turban et al. 2011, p. 361.) Furthermore, also for instance Hackathorn (2004, p. 4) states that a more timely data is needed because in some cases the data is more valuable when it is fresh. He however argues that real-time is misleading term and that companies should consider the real value to business. This means that the data should be timely for the case in hand, not necessary real-time.

2.3.3. *Repositories*

When data is used in analytics, it is structured in a certain way into a repository which is specially designed to organize and store analytical data (Davenport & Harris 2007, p. 166). These repositories are called usually either data warehouses or data marts. In addition also metadata and metadata repository is part of the storage of analytical data.

Data warehouse and data mart

In short, data warehouse is a database, which contains regularly updated data from multiple different data sources. It contains large amounts of historical data, which can be used to facilitate different kinds of analysis over time (Davenport & Harris 2007, p. 166). Turban et al. (2008, p. 39) state that a data warehouse is a pool of data which is structured in an optimal way for analytical processing activities such as data mining, OLAP, and reporting. Ideally a data warehouse is totally integrated and it supports the needs of an entire company (Turban et al. 2008, p. 39). Such a data warehouse is usually referred as an enterprise data warehouse.

Turban et al. (2008, p. 39-40) define that a data warehouse is subject oriented, integrated, time variant, and nonvolatile. This means that the data is organized by a certain subject, for example customers. Data warehouse also integrates data from multiple data sources into one common consistent format. Therefore the data warehouse provides a more comprehensive view than operational databases which are usually product oriented and tuned for handling transactions. (Turban et al. 2008, p. 39). In addition, data warehouse is used as a system of record for reporting and analysis. This means that the data warehouse is the authoritative data source for the data elements which are stored in the data warehouse. Additionally, as stated, Turban et al. (2008, p. 40) define that time is also one important dimension of a data warehouse. Data warehouse must support data from multiple different source systems which all can contain multiple different time points such as daily and monthly views. On top of historical data, real time data warehouses provide also current data. Furthermore, Turban et al. (2008, p. 40) also argue that traditional data warehouses are nonvolatile, meaning that users cannot change the data in the data warehouse. For this reason the data warehouse can be designed to support fast data access instead of fast data update.

Data mart on the other hand is a smaller database than a data warehouse and it is usually related to a single business function or to a certain process (Davenport & Harris 2007, p. 166). Therefore it is a subset of a data warehouse and usually contains data for a single subject area such as marketing (Turban et al. 2008, p. 40). Data mart can be a separate database or it can be a partitioned section of the central data warehouse (Davenport & Harris 2007, p. 166). Turban et al. (2008, p. 40) define that there are two types of data marts; dependent and independent. Dependent data mart gets its data from the data warehouse thus using the same consistent data model and quality data as the data warehouse. This kind of approach ensures that all users, even users of different data marts, access the same version of the data. Independent data mart on the other hand is not connected to a data warehouse. It can be described as a small data warehouse which is only designed for a certain department or function. (Turban et al. 2008, p. 40).

Metadata repository

As discussed earlier, metadata is information about the business data. Davenport & Harris (2007, p. 166) state that this data definition includes information about the data source, unit of measurement, how the data has been calculated, and bibliographic information. Instructions about how the data should be used and information about data reliability and accuracy are also attributes that may be included in metadata. Turban et al. (2008, p. 41) define that metadata can be categorized into three categories. Syntactic metadata describes the syntax of the data, structural metadata describes the structure of the data, and semantic metadata describes the meaning of the data in a specific domain.

As a further definition, Turban et al. (2008, p. 41) argue that metadata provides enriching information about the context of the data and therefore assists in the conversion of data and information into knowledge. Thus metadata strongly contributes to the effectiveness of data use and is a critical part of data warehousing. Davenport & Harris (2007, p. 166) state that metadata should be stored in a common metadata repository which is used by all analytical applications in order to ensures that the data is consistent.

Different viewpoints on data warehouse architecture

There exists multiple different ways to build a repository architecture for analytical data. These approaches include active data warehouse, federated data warehouse, and data mart data warehouse among many others. One widely accepted and currently widely used approach is an enterprise data warehouse. (Inmon et al. 2008, p. 12; Turban et al. 2008, p. 43-44; Eckerson 2010). This approach to data warehouse architecture is illustrated roughly in Figure 8.

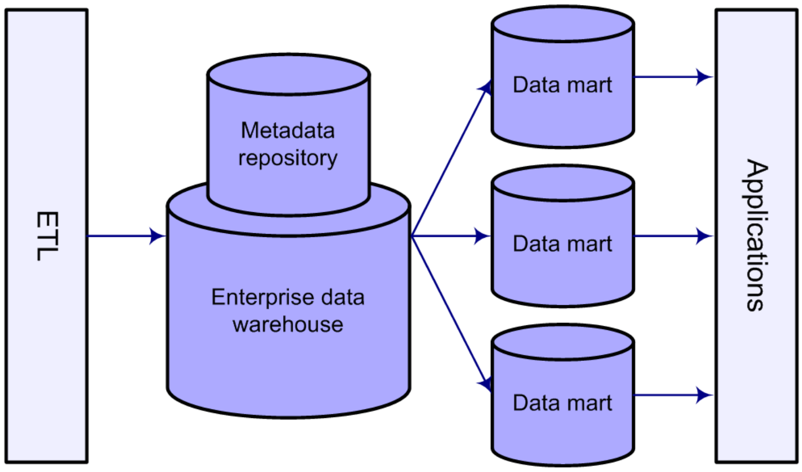

Figure 8. A possible data warehouse architecture (partly adapted from Turban et al. 2008, p. 43).

The described enterprise data warehouse approach has multiple different variants which are used in different companies but high level specifications can be made. Turban et al. (2008, p. 41-44) define that in this approach the enterprise data warehouse is fed with atomic-level, cleansed and organized data from source systems across the enterprise through ETL process. The large-scale nature of enterprise data warehouse allows integration of data from multiple sources into one common standard format, which can be used as a basis for effective business intelligence and related activities.

After integration the data is either first summarized or forwarded as such into different data marts (Eckerson 2010). Each data mart contains data from a specific area and therefore they do not have such vast amounts of data as the enterprise data warehouse. Metadata is stored centrally and it is exchanged throughout the data warehouse architecture (Turban et al. 2008, p. 43). With their subject oriented data, data marts serve certain functions or departments in their business intelligence and analytics related activities.

The smaller data amounts of data marts allow faster analysis when compared to the enterprise data warehouse. If an enterprise-level view is needed for business intelligence and analytics related activities, the enterprise data warehouse can be used to get additional data. Data warehouse can also provide atomic-level data for the business intelligence and analytics related activities if data mart only stores summarized information.

2.3.4. *Analytical tools and applications*

As discussed in previous sections different analytical tools and applications are an important part of business intelligence architecture especially when business analytics is concerned. In the coming sections main analytical tools and applications, such as data mining and text mining, are introduced. Especially data mining is described in detail as it can be used in multiple different application areas of business analytics. In addition, the different application areas of business analytics which are related to the customers are discussed more in chapter 3.

OLAP

OLAP tools are used to analyze multidimensional data. Davenport & Harris (2007, p. 168) state that these tools are able to analyze data models which have even seven or more different dimensions such as product line, time, geography, and so on. Furthermore, Thierauf (2001, p. 113) writes that especially such dimensions are used which are important and used to manage the business. Because of their ability to analyze multidimensional data these tools are especially used for semi-structured decisions and analyses (Davenport & Harris 2007, p. 168). Additionally, for instance Turban et al. (2005, p. 260) state that OLAP tools include such features as dimension expansion and collapse, automatic calculations, instantaneous drilldown and rollup, and so on. This means that users can for example use these tools to drilldown on a specific dimension of data. One example of such case could be drilldown from the product line level summary data into the detailed data of all the products which belong to that product line.

Data mining

Data mining tools use multiple different techniques to extract hidden, descriptive and predictive information and to indentify patterns even from complex and poorly defined databases. As an example these tools can be used to predict which customers are likely to churn[1]. (Davenport & Harris 2007, p. 169; Turban et al. 2008, p. 103.) When defining the difference of data mining and OLAP, one can look at the questions the different technologies can answer. OLAP can provide the answers for many of the questions that users need to be solved but data mining on the other hand can also answer questions users do not even understand to ask (Turban et al. 2008, p. 103).

In practice data mining tools are used to create different models which aim at solving business problems. These models describe patterns and relationships found from the data. (Rygielski et al. 2002, p. 487-488.) Ngai et al. (2009, p. 2593) write that there exist seven different types of models which cover the generally used data mining models. These include association, classification, clustering, forecasting, regression, sequence discovery, and visualization. Rygielski et al. (2002, p. 488) state that association and sequence discovery are used to describe behavior, classification and regression are used to make predictions, and clustering can be used to do both. Forecasting is used to estimate the future value of the record based on the patterns found from the data and visualization on the other hand is used to present complex patterns visually (Ngai et al. 2009, p. 2595). Furthermore, each type of data mining model can use multiple different algorithms to process the data. Ngai et al. (2009, p. 2593) write that the commonly used algorithms include association rule, decision tree, genetic algorithm, neural networks, K-nearest neighbour, and linear or logistic regression. Additionally, for instance especially association rules, decision trees, neural networks, and linear and logistic regressions are used in predictive modeling (Zaman 2005, p. 1; Eckerson 2007b, p. 7).

The process of data mining can be generally described to include six phases. These phases are illustrated in Figure 9 and shortly described in the following paragraphs. In business understanding phase the objectives and requirements for data mining are defined together with the business. The main goal is to define what is the problem or question that is wanted to be answered. The second phase, data understanding, includes data collection and the process of getting to know the data. The focus is to determine possible data quality problems and to identify data that might be interesting concerning the problem or question in hand. (Chapman et al. 2000, p. 13-14).

In data preparation the actual tables, records, and attributes, which will be further processed and inserted into the modeling tools, are selected. Also the possibly needed data transformation and cleaning is done in this phase. The fourth phase, modeling, includes the selection of used modeling techniques and the optimization of the parameters these techniques use. (Chapman et al. 2000, p. 13-14).

[1] Customer churn and churn prediction are widely used terms in research and in business. Churn prediction includes identifying those customers which are likely to leave the company in near future and move doing business with company’s competitors.

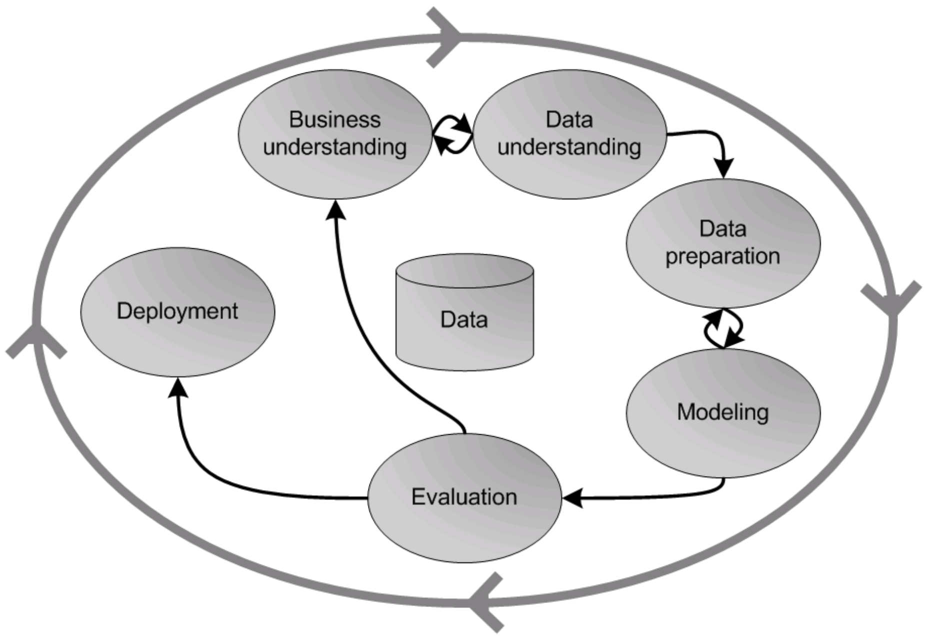

Figure 9. Phases of data mining project (adapted from Chapman et al. 2000, p. 13).

In the evaluation phase on the other hand the validity of created models is evaluated. The main focus is to check that all important business issues are considered in the model. If the model is the answer for the business problem or question, it is deployed in the last phase. The actual steps or tasks included in deployment vary a lot between different projects. The models might be applied to the decision making processes of an organization or a simple report might be created based on the model. (Chapman et al. 2000, p. 13-14).

Text and web mining

Text mining tools and applications apply data mining to unstructured text documents. These tools help to find hidden content from documents, connect documents in to previously unknown categories, and group documents by common themes. The basic functionality behind text mining is to map unstructured data into a structured format. This is done by calculating the frequencies of different terms in a text. (Turban et al. 2008, p. 159-161.)

Text mining tools can also for example identify and count words on different kind of websites such as blogs, forums, and so on in order to identify new trends or relationships (Davenport & Harris 2007, p. 170). These tools are also called as web mining tools (Turban et al. 2008, p. 163). On top of the capabilities of text mining tools, web mining tools can not only develop useful information, such as popularity of a certain document, from the links included in web documents, but also extract useful information from web page visits and transactions (Raisinghani 2004, p. 149-156; Turban et al. 2008, p. 163-165). Web mining is one of the promising application areas of data mining as Internet is a dynamic source of information (Raisinghani 2004, p. 149).

Rule engines & simulation

Rule engines are applications that use conditional statements to describe business rules. These business rules can therefore address logical questions and provide recommendations to decision makers. (Davenport & Harris 2007, p. 169.) Rule engines can also be used as a part of larger automated applications. When an event that is associated with business rules happens, rule engine evaluates the event against its stored business rules and determines what action needs to be taken (Turban et al. 2008, p. 115). This enables companies to adapt and respond dynamically as the business rule management is centralized in robust repositories. Additionally, there exist optimization applications which find mathematically the best solution to a complex problem with many decision options and constraints. Davenport & Harris (2007, p. 58) write that optimization applications can find the best solution to particular objectives by efficiently allocating limited resources taking in to consideration existing restrictions.

Furthermore, also specific applications to model business processes exist. By using these simulation tools users can simulate the change of a business process or streamline the flow of information or products. (Davenport & Harris 2007, p. 170.) LaValle et al. (2010, p. 13) define that simulations technologies are able to recommend automatically optimal approaches by evaluating alternative scenarios. An example from such a simulation could be the ideal number of sales representatives for a new territory.

Emerging analytical technologies

Additionally to the already mentioned analytical applications and tools Davenport & Harris (2007, p. 170-171) mention also multiple other emerging analytical technologies. These include text categorization, genetic algorithms, expert systems, audio and video mining, swarm intelligence, and information extraction.

Text categorization tools rate document’s relevance to defined topic by using statistical models or rules. Genetic algorithms on the other hand optimize random processes such as delivery routes. (Davenport & Harris 2007, p. 170.) Furthermore, expert systems are specialized artificial intelligence applications which advice decision makers and audio and video mining tools look patterns from full-motion images and sound. Additionally, swarm intelligence tools help users to understand what kind of effects low-level changes to a system can have and information extraction finds specific concepts from large amounts of textual data. (Davenport & Harris 2007, p. 171.)

2.3.5. *Presentation tools and applications*

Presentation tools and applications such as different reporting tools or portals provide access and visualization to the data in the data warehouse. These tools and applications can also be used to deliver different reports, scorecards, alerts and such to end users. From the business analytics perspective presentation tools and applications can be used to deliver the insight acquired through the use of analytical tools and applications to the correct persons in a company. (Davenport & Harris 2007, p. 171). For example reports, dashboards, and scorecards can be used for broad distribution of different analyses throughout the enterprise. Also LaValle (2009, p. 7) states that for instance predictive dashboards and visualization tools make it easy to grasp the meaning of information. This is critical as the insight is only valuable when it results into a worthwhile action.

Furthermore, Davenport & Harris (2007, p. 172) argue that proper presentation tools and applications are truly important when a company wants to leverage analytics by allowing also users without any statistical skills to manipulate data and execute analyses. Certain visual analytical tools have an intuitive visual interface, which allows the creation of sophisticated analyses without a need to modify underlying statistical models, thus making the usage easy for anyone. (Davenport & Harris 2007, p. 172).

3. CUSTOMERS AND BUSINESS ANALYTICS

In this chapter one of the main concepts of this study – customer analytics – is defined in detail. Firstly the different types of customer data and the different concepts which are related to customer analytics are discussed. Then the customer analytics itself is described both in high level and more in detail by introducing different application areas of customer analytics.

3.1. Customer information

Companies can have multiple different types of customer data from both individual and business customers. Some of that data can be linked to an identified customer and some of that data can comprise a larger group of unidentified customers (Arantola 2006, p. 52). Additionally, Arantola (2006, p. 52) states that customer data can also be obtained through multiple different sources or methods such as internal operational systems and customer interviews or researches. In high level customer data can be categorized into three main categories which are basic customer data, behavioural data, and predictive data.

Basic customer data which describes customers includes demographic data, geographical data, socio-economic data, psychographic data, and general behaviour and attitude related data. Demographic data encompasses different fundamental attributes of a specific customer. These attributes include age, gender, marital status, number of children, and so on. (Stone et al. 2004, p. 113.) Geographical data includes such attributes as street address, zip code, and country. Stone et al. (2004, p. 113) state that also customers relative location to other customers with certain socio-economic or demographic attributes can be useful. Furthermore, Arantola (2006, p. 52) argues that basic customer data is usually collected routinely. Good examples of such routines are loyalty customer card application form or order form where contact information is collected. After the forms are filled in, the data can be inserted into company’s internal systems. Demographic data can also be collected from different external sources, such as public customer data sources like population register or company register (Arantola 2006, p. 72).

Socio-economic data, as another aspect of the basic customer data, includes details such as the occupation, income, and assets of a specific customer. This data is often seen as valuable as it indicates buying power and aspirations. (Stone et al. 2004, p. 113.) Psychographic data on the other hand defines customer’s attitudes, interests and opinions such as is the customer extrovert or introvert, or is he consumer or saver. This kind of data is usually collected with different customer questionnaires. (Stone et al. 2004, p. 114.) Therefore, in many cases psychographic data cannot be connected to an individual customer but it is instead connected to a specific customer group or segment. The same situation applies for data related to general behaviour and attitudes of the customers. Stone et al. (2004, p. 113) write that an example of such data is for instance the shopping preferences of a specific customer; if one enjoys shopping or not.

Investigating the categories of customer data further, the behavioural data, telling how a customer behaves, includes transactional data, interaction data, and data related to actual behaviour and attitudes. Stone et al. (2004, p. 36) write that transactional data includes information on who bought or returned a specific product, when the transaction occurred, and how it occurred. It can also include additional detailed data such as the price of the product, related promotion, and so on. And although transactional data can be related to a specific customer, Arantola (2006, p. 74) argues that it is possible that a company is not able to uniquely identify the related customer. Instead, the transactional data might be connected to a group of customers or to a specific location or product. Transactional data can be collected through company's transactional systems. Stone et al. (2004, p. 36) state that transactional data has to include a sufficient level of detail as it is one of the most important indicators of likely future transactions. They also argue that based on years of experience frequency, recency, amount and category related transactional variables dominate the most explanations of buying behaviour (Stone et al. 2004, p. 36).

Customer data related to actual behaviour and attitudes comprises aspects like whether or not a customer is a user of a specific product or category, frequency of use, loyalty, and so on (Stone et al. 2004, p. 113). This data is closely related to transactional data as transactional data can be used to define behavioural data such as purchasing behaviour. On the other hand, interaction data, as the last aspect of behavioural customer data, considers all the remaining behavioural data company can collect from its customers. For instance online interaction data such as internet click-stream data tells much about the customers (Fayyad 2003, p. 2). Interaction data can also include for example emails, telephone calls and business replay cards and it can be gathered from different customer touchpoints (Hauser 2007, p. 40).

Additionally, predictive customer data predicts future customer behaviour. It can be acquired for example by making statistical predictions from the other customer related data; basic customer data and especially behavioural data. These statistical predictions can be made by using different predictive models such as customer churn prediction model or best offer model. (Arantola 2006, p. 72.) Especially predictive customer data is discussed more in detail in the upcoming sections.

Finally, when customer data is being discussed, it is always important to remember the data privacy related subjects. Rygielski et al. (2002, p. 495) argue that when companies analyze customer data, they balance between building better relationship with their customers and the risk of invading customer privacy. The analysis of customer data can bring insight to be used to help the company to serve its customers better. On the other hand company might learn such things from its customers, which the customers do not feel comfortable with letting the company know about. For this reason it is important that companies find a balance between privacy rights for customer protection and still be able to provide benefits to businesses (Rygielski et al. 2002, p. 495). Also Davenport & Harris (2007b, p. 8) argue that companies which analyze customer data need to keep "win-win" in mind. This means that both the customers and the company should benefit from the analysis of customer data. Furthermore, there also exists different data privacy related legislations in different countries which need to be followed. All in all, as Stone et al. (2004, p. 219-223) state there are many different social, political, and legal issues which companies need to consider when using customer data.

3.1.1. *Single view of customer*

As stated, nowadays companies can have collected large amounts of data about their customers. However, even though a company has a much used customer database with excellent quality customer data, it does not necessarily yet mean that the company really understands its customers. Liautaud & Hammond (2001, p. 138) argue that it requires much more than just knowing demographic distribution of your customer base to fully understand your customers. Examples of the other aspects needed for understanding your customers are the multiple different types of transactions and interactions conducted in sales channels and customer touchpoints. Additionally, it is also important to understand that there can be many different sales channels and customer touchpoints in a specific company. Sales channels can include for example different direct sales channels, online channels, reseller based channels, and so on. Customer touchpoints can include for instance marketing programs, customer support centres, sales force interactions, website, email, and so on.

If a holistic picture of the customer is formed from all this information and from the common customer related demographic information, a 360° view of the customer can be acquired (Liautaud & Hammond 2001, p. 138). Furthermore, for example Raisinghani (2004, p. 267) writes that those companies which integrate customer data from large number of different internal and external data sources into enterprise data warehouse are able to understand their customers and their relationships with them better. Additionally, Raisinghani argues that this understanding can be used to achieve greater profitability when compared to other companies. Single and consolidated view of customer creates visibility into customer's total portfolio value and into households and organizational structures. This information can then be used in business processes and applications throughout the organization. On top of that, for example Stone et al. (2004, p. 143) argue that also the customers should believe that they are making business with one integrated and complete company rather than with a disjointed set of business units.

Liautaud & Hammond (2001, p. 139) state that it is because of multiple and disparate operational systems, that many companies have not acquired a 360° view of their customers. Different departments and business units often have their own systems and possible mergers and acquisitions even increase the amount of different systems. All these systems should be integrated so that company could have a common customer data. For example Schroeck (2001, p. 2) states that the lack of integration between company's different systems limits the usefulness of each individual system. Another degree of complexity is introduced when data is inconsistent. If for instance a sales system and customer service system both have a customer in their database, it does not automatically mean that they can be easily connected (Liautaud & Hammond 2001, p. 139). The customer can have for example totally different customer number and even the customer name could be mistakenly typed differently in the other system.

3.1.2. *Customer insight*

Term *customer insight* has often been used as a synonym for customer data but also other definitions exist. Arantola (2006, p. 53) states that customer data and customer insight are not synonyms and thus introduces different definitions used in literature for customer insight. Customer insight might mean deep understanding of customer behavior and needs or an approach where customer has a centric role as a producer of understanding. It can also refer to a company's function which collects customer related information into one common place, or that company is able to collect all information related to a specific customer from its different business units or departments. (Arantola 2006, p. 53.) Arantola (2006, p. 152) itself argues that customer insight means enriched customer information which is connected to the use case in hand and therefore can be used in decision making. In this research customer insight is defined as a deep understanding of customer behavior and needs which can be used to make decisions. This understanding is derived from customer information which is collected from across the organization.

Customer insight helps companies to know their customers better by knowing among other things their product and service needs, purchasing habits and behavior, characteristics, loyalty, profitability, and interaction preferences (Schroeck 2001; Turban et al. 2005; Arantola 2006; Turban et al. 2008). This can be used for achieving correct and timely actions, increased revenue per customer, customer loyalty, new customers, just to name a few.

3.2. Related concepts

There are many concepts which consider companies and their customers. In order to understand where customer analytics fits among these other concepts, it is beneficial to shortly describe a few of the related concepts. The two most closely related concepts to customer analytics are customer relationship management and customer intelligence.

3.2.1. *Customer relationship management*

When customer analytics and customer data is discussed, it is purposeful to first define the basic concept of Customer Relationship Management (CRM). CRM is not just a matter of how a company works with its customers but it also defines how a company solves problems for its customers, encourages its customers to buy products or services, and deals with the financial transactions. One can also say that CRM is a business strategy and therefore much more than just a system. CRM system however is often an important part of CRM. For example Voudouris et al. (2008, p. 205) state that many companies have introduced CRM systems to contain valuable customer information which can be used for improving company's customer relationships and services. Additionally, Cunningham (2002, p. 6, 112) defines that CRM is a system that includes all the aspects of company's interactions with its customers. CRM also improves and optimizes the methods used for managing the interface and interactions between company and its existing and prospective customers.

Supporting the previous claim on CRM's versatility, Cunningham (2002, p. 32) states that CRM has been usually mistakenly viewed as a sales system and customer management as a sales issue. CRM however includes multiple different aspects such as sales, marketing, support, finance, contracts, operations, and so on. In many of the CRM related studies the following four dimensions of CRM are identified: Customer identification, customer attraction, customer retention, and customer development (Ngai et al. 2009, p. 2593).

The CRM process starts with customer identification where that part of the population, which would be the most profitable for the company or the most likely to become customer, is identified. In customer attraction phase these identified customers are targeted with different marketing activities. Following the attraction phase, in the customer retention phase customer loyalty and satisfaction and other aspects, related to long term customer relationships, are focused on. At the same time in the customer development phase such facets as customer profitability and up- or cross-sell possibilities are kept in mind. (Ngai et al. 2009, p. 2594-2595.)

In an optimal situation a CRM system stores all the relevant customer information which was also discussed in section 3.1. CRM also defines and enforces the processes related to the different dimensions of CRM. That said, a CRM system does not necessarily have any analytical features. For example Arantola (2006, p. 135) states that a study containing some of the Finnish top 500 largest companies revealed that most of the companies used operative CRM systems to handle customer relationships and to store customer related information. However, only few of these companies had tools and applications which would help in the analysis of customer information; in analytical CRM.

Operative CRM provides the data and processes for the user who then does the necessarily analyses and acts based on those analyses. Such analyses are usually descriptive in nature (Arantola 2006, p. 136). A good example of this is for example segmentation. A user might decide to create buying behavior related segmentation into the CRM system based on to the country of a customer because he thinks that it is an attribute which can be used to group customers with similar buying behavior. This decision is based on to the analyses the user makes based on to the data in CRM system and to his previous experience. However, analytical technologies make it possible to identify the actual attributes which can be used to group similar customers (Stone et al. 2004, p. 155). Such technologies also make it possible to predict customer behavior (Arantola 2006, p. 136). In this research analytical CRM is seen as part of customer analytics which is discussed in the upcoming sections.

3.2.2. *Customer intelligence*

In addition to customer relationship management also customer intelligence is closely related to customer analytics. The difference between these terms is not however defined here in much detail as the differences match closely the differences of business intelligence and business analytics which were discussed in section 2.2.1. In short customer intelligence focuses on to the customer data access and customer related reporting. This data access and reporting can be used for instance to support sales, marketing and service related functions. Such reports can include for instance sales reports or service activity reports for different customer segments. Similar as with business intelligence the basic idea of customer intelligence can be seen as the transformation of customer data into customer information which can be used to make smarter decisions resulting to better actions.

3.3. Customer analytics

As shortly already discussed regarding customer relationship management, different analytical technologies enable among other things to group similar customers and predict customer behavior. This is important because today's customer value-drivers and demographics are changing and responding to these changes is critical for companies targeting for market leadership. Many outdated business models need to be transformed to meet the new customer demands. (Gonzalez-Wertz 2009, p. 5.) Also the amount of data, companies have collected about their customers, has exploded in the last few years. Now the challenge is to find those bits of information among the countless data that add real value to decision making and enable companies to respond to the changing customer demands. This can be achieved by creating customer insight mining the data and leveraging analytics. (Gonzalez-Wertz 2009, p. 2-3.) The top performers derive value from customer insight and use it across all customer touchpoints, use new data types and data sources, and deliver consistent and better experience to their customers (Gonzalez-Wertz 2009, p. 8).

Companies have analyzed their customers for a long time through which there are also multiple different concepts related to the analysis of customer data. For example Thierauf (2001, p. 246) writes that companies, which focus on competitive advantage, practice database marketing. Database marketing is defined in this context as a set of activities including for instance the collection, analysis, and use of individual customer attributes and behavior patterns in supporting different areas from promotional selling to strategic market analysis. On the other hand, Schroeck (2001, p. 4) defines another term, customer analytics, as application of advanced analytical methods to the customer data of a company in order to help business in identifying, attracting, and retaining the best and most profitable customers. Furthermore, Schroeck states that customer analytics makes it also possible for instance to anticipate, measure, and influence customer behavior. In addition to marketing this definition includes the application of customer understanding also to the sales and service related processes. Similar definitions for customer analytics can be found also from other researchers (Lenzen 2004; Davenport & Harris 2007b; Voudouris et al. 2008). In this research customer analytics is defined as the application of advanced analytical methods to the customer data of a company in order create customer insight.

Understanding the definition of customer analytics, one can see that customer analytics and the created customer insight can be used for supporting all the dimensions of the customer relationship management, including but not limited to customer identification, customer attraction, customer retention, and customer development. Ngai et al. (2009, p. 2594) write that in literature for example customer segmentation, target customer analysis, one-to-one marketing, up-selling, cross-selling, market basket analysis, and customer lifetime value have been defined as data analysis techniques which can be used for supporting different dimensions or phases of CRM.

Additionally, especially both explanatory and predictive analytics are used in customer analytics. Explanatory modeling is used for creating new understanding about customers, whereas predictive modeling is used for predicting customer events and needs. Creating new understanding about customers helps organizations in making information based decisions. All industries can take advantage for example from customer segmentation created with data mining (Rygielski et al. 2002, p. 490). Also predicting what customers want and how they will react is important (Voudouris et al. 2008, p. 205). Predictive analytics allows companies to synthesize different customer data in order to be able to act at the point of impact. For instance when companies are faced with customer opportunities or challenges, predictive analytics can help them to determine causal factors and alert companies to take appropriate actions. This is really beneficial as for example Stone et al. (2004, p. 160) write that customer retention and customer development are important things for companies as they both need to be increased in order to be able to compete with other companies. A good example of this is telecommunication industry where the competition is really hard. Voudouris et al. (2008, p. 207) write that because of the characteristics of telecommunication industry it is especially important for telecommunication companies to be able to predict which customers are likely to churn, who are under stress due to failed processes or who are correct targets for cross-selling.

Furthermore, different predictive models can for example predict the probability for a specific customer event or even estimate the time when that customer event occurs (Voudouris et al. 2008, p. 207). Zaman (2005, p. 2) for example tells that credit card companies can predict the loss of a customer by creating predictive models based on frequency of use and personal financial situation. In addition the *annual percentage rates*[1] offered by different competitors can be included to the model. This predictive model can be then applied to those customers who start using their credit cards less frequently. Predictive analytics would group these users, find the pattern of card usage for this group, and predict the probable outcomes.

In the next sections some of the different application areas of customer analytics which have been discussed in other researches are introduced. The application areas are grouped with the dimensions of CRM which were described when the concept of CRM was defined.

3.3.1. *Customer identification & attraction*

The application areas of customer analytic which are related to customer identification and customer attraction focus on creating understanding about the basic structures of the customer base and to target these customers. Ngai et al. (2009, p. 2594) state that customer identification includes such concepts as customer segmentation and target customer analysis. Target customer analysis includes the identification of customer profitability enabling the targeting of the most profitable customer groups (Woo, Bae & Park 2005, according to Ngai et al. 2009, p. 2594). Furthermore, Ngai et al. (2009, p. 2594) state that customer attraction includes the selection of those customer segments which will be targeted.

Sophisticated segmentation

Arantola (2006, p. 85) writes that sales and marketing are functions where the need to understand customers better usually arises in companies. One method which is used for achieving this understanding is segmentation. Stone et al. (2004, p. 112) state that segmentation is the process of grouping similar, but not identical, customers. There are multiple reasons behind identifying such segments and one of them is their usability in predicting the behavior of customers more accurately (Stone et al. 2004, p. 112). When

---

[1] An annual percentage rate (APR) is a standardized method of quoting the effective interest rate on consumer loans. It is used especially when interest is computed on monthly basis. An APR includes all fees, and takes into account the continual reduction of principal amount through amortization.

a company knows how other similar customers behaved, it can predict especially with the help of analytical applications how a specific customer will behave. Traditionally segmentation has been for example used for identifying target groups for marketing, or for identifying the most valuable customers for the effective targeting of the sales activities. Moreover, segmentation can be used for identifying customers with similar lifestyle or consuming habits as a basis for new product and service development (Arantola 2006, p. 85). Simple segmentation of customers has usually been created by using basic customer data such as age or location as a basis for the segmentation. But because these traditional segments are large and include still high variability, such segmentation typically ends up telling us quite little about the individual customers of these groups. (Stone et al. 2004, p. 155.)

Different analytical tools and applications with data mining functionalities can be used to identify more detailed or sophisticated customer segments. Stone et al. (2004, p. 155) state that by using data mining companies can define customer classes, segments, and assign individual customers to these classes by using a specified criteria. This kind of an approach uses classification as a data mining model (Ngai et al. 2009, p. 2595). In marketing for example such segmentation is often used for designing offers more easily to specific customers and to target them more precisely (Stone et al. 2004, p. 112). In many cases segmentation also includes potential customers. Companies can therefore use the segments of potential customers to focus the efforts and resources in order to attract these potential customers (Ngai et al. 2009, p. 2594). It is also important to remember that segments are not necessarily static. Some segments can change over time and customers can move between different segments (Stone et al. 2004, p. 112). Different analytical tools and applications can be used for redefining customer segments often enough so that such changes are identified.

Another way of grouping similar items into segments is cluster analysis. Stone et al. (2004, p. 155) write that cluster analysis uses a numeric criterion of similarity or proximity as a basis for the grouping. Cluster analysis does not use predefined groups or important factors, instead the groups and important factors are chosen so that they minimize the differences within each group and maximize the difference between the groups. Other segmentation tools include for example neural networks, decision trees, discrimination analysis, and genetic algorithms (Chan 2008, p. 2755; Ngai et al. 2009, p. 2595).

The variables or important factors used in segmentation can vary a lot. All demographic, psychographic, and socio-economic variables as well as purchasing or consuming behaviour can be used when customers are segmented (Stone et al. 2004, p. 113-116). On top of that Chan (2008, p. 2756) argues that for example in customer targeting it is important to include also customer profitability into the segmentation model. All in all, segmentation is an important part of many application areas of customer analytics. These application areas include for example Market Basket

Analysis where the point of sale data can be combined with customer information in order to identify correct product offers for specific customer segments.

Besides segmenting or classifying the identified customers, companies can also create customer profiles from the data which they cannot connect into identified customers. As these profiles usually consist of customer attitudes, opinions and such aspects of data which can be collected with market researches, they can contain data from both active customers and competitor's customers. These profiles can for example define what type of customers there are in a specific area. Additionally, also the company's website users can be profiled even though they are not identified during the session. This can be done by analyzing the web surfing behavior of the users with click-stream analysis.

Customer profitability

On top of understanding customers' or customer segments' needs and wants, it is also beneficial to define the relative profitability of individual customers. Relative profitability can include actual and potential profitability as well as current and lifetime profitability. (Stone et al. 2004, p. 125.) Current actual profitability of a customer can be calculated if all the costs and profits related to that specific customer are available (Arantola 2006, p. 151). This requires a comprehensive view to different IT systems handling customer information so that a full picture about the actual profitability can be formed (Arantola 2006, p. 87). However, Arantola (2006, p. 97) states that companies in consumer market often evaluate the value of a certain customer only based on the socio-economic data such as income and additionally companies in business to business market often evaluate the value of a certain customer only based on its size.

Current potential profitability for potential customers on the other hand can be determined by defining the current sales potential for that specific potential customer. Same applies for up-selling and cross-selling possibilities for a specific existing customer. For example Hwang et al. (2004, p. 185) state that different data mining techniques such as decision tree, artificial neural network, and logistic regression can be used for achieving this. Firstly, the variables affecting to the fact whether customers use a product or service or not are derived by analyzing the whole customer base after which these variables are compared to specific customers. Both actual and potential lifetime profitability can be estimated by using for example different customer lifetime value models (Hwang et al. 2004, p. 182). Customer lifetime value and its application areas are discussed more in detail in section 3.3.3.

Customer profitability can also be included into the customer segmentation model discussed in the previous section. For example Chan (2008, p. 2756) argues that for customer targeting and marketing campaign management it is important to also include customer profitability into the segmentation model. Such segmentation models can be for example used for focusing marketing campaigns for those potential customers segments which have the highest potential profitability. This way the customer attraction can be focused on to the most valuable customers from company's perspective. Also Stone et al. (2004, p. 125) argue that an increasing number of companies use customer profitability in segmentation. They however state that measuring or estimating customer profitability can be problematic. In some cases customer profitability can be calculated to some extent from the weight of purchasing or weight of use (Stone et al. 2004, p. 126). Similar approach as analyzing customer profitability can be used also to minimize risk for example by banks. Hsieh (2004, p. 632) writes that banks can use data mining techniques such as neural networks and association rule inducers to score new credit card applicants based on the available customer data and consequently use the scoring when deciding whether or not to grant credit.

3.3.2. *Customer retention*

The application areas of customer analytic which are related to customer retention focus on keeping the customers loyal. This includes for instance understanding customer's expectations, level of satisfaction or possibility to churn. The goal of the application areas related to customer retention is to provide such insight which can be used for maintaining a long term relationships with customers. (Ngai et al. 2009, p. 2595.)

Customer loyalty

As stated, another important application area of customer analytics is related to customer value, loyalty, and satisfaction. Voudouris et al. (2008, p. 208) write that both customer satisfaction and customer loyalty are important for future improvements and therefore many companies conduct for example different customer satisfaction surveys. These surveys are done because companies want to understand their customer's attitudes, expectations, and loyalty more in depth. They also want to understand what are the drivers and internal processes which affect customer satisfaction and loyalty. (Voudouris et al. 2008, p. 208.) With this information and by using analytical applications it is possible to identify the degree of loyalty of a specific customer. Additionally, the degree of loyalty a customer exhibits can further on be used in for example focus marketing (Stone et al. 2004, p. 126). Loyal customers might be targeted with different kind of marketing and sales activities than customers with lower degree of loyalty.

The drivers of customer loyalty definitely need to be used for increasing customer loyalty. If a company knows what drives different customers to continuously make new purchases, it can develop those elements further. This gives companies a good opportunity to increase sales by increasing the amount of loyal customers. However, Stone et al. (2004, p. 126) argue that measuring loyalty is not a simple task and that in some cases it is not even reasonable. They also write that there exist multiple behavioral and attitudinal indicators of commitment or loyalty. On top of that these indicators might vary between different customer types. Also Voudouris et al. (2008, p. 208) write that, as the drivers of satisfaction typically are not linearly cross-correlated, a simple linear statistics is not enough in order to analyze drivers of satisfaction but instead the analysis requires a more sophisticated approach. Stone et al. (2004, p. 127) however state that there exists different statistical methods, which can be used for inferring the factors leading to a good customer retention level.

Churn prediction & customer events

On top of understanding what drives satisfaction or value for the customer, it is also important to understand the different drivers behind customer churn. With data mining and predictive analysis tools and applications companies can analyze the customer data and metrics in order to gain insight that helps them to find churn candidates and to target them with marketing campaigns. Timely and attractive promotions may sway unsatisfied customers away from switching to a competitor and thus provides a way for companies to keep their clients loyal (Zaman 2005, p. 2). This gives companies an enormous opportunity for revenue assurance through customer retention.

The implementation of customer churn prediction can be illustrated with an example related to e-commerce environment. Boyer et al. (2005, p. 570) write that by combining subjective customer perceptions of their online ordering experience from surveys, with the actual transactional data related to these customers, companies can predict which customers are unlikely or likely to repurchase and which customers are undetermined. This is achieved with the help of data mining tools and predictive models. These predictions are extremely useful for operational purposes and they can be used for example in targeting the marketing activities so that those customers, who could be persuaded to repurchase, are targeted with special offers (Boyer et al. 2005, p. 570). Therefore these predictions also help companies to increase revenue through sales and marketing as less effort is needed to compensate the customer churn.

As another example, Schroeck (2001, p. 4) argues that in telecommunications industry customer analytics can be used to help determine both the customer lifetime value and the likelihood of attrition over certain period of time. This information can then be used for focusing retention programs so that those customers, which have the high possibility to churn and high value to the company, can be targeted. When the churn prediction is combined with the information about customer lifetime value like this, even greater business value can be acquired comparing to the basic churn prediction.

Companies can analyze also other customer related events besides customer churn. These events can include such as change in address, birth of a child, or change in marital status. By analyzing these customer events companies can identify the specific events which require reaction in a way or another. For example birth of a child could be an opportunity to increase sales, whereas change in address could be seen as a danger to loose the customer, and so on.

3.3.3. *Customer development*

The application areas of customer analytic which are related to customer development focus on consistent expansion of transaction intensity, transaction value, and individual customer profitability. These application areas include for instance customer lifetime value analysis, up-selling, cross-selling and market basket analysis. (Ngai et al. 2009, p. 2595.) Stone et al. (2004, p. 160) argue that the main focus for many companies is to sell more to customers through cross-selling, more often by increasing selling frequency, and for a longer time. This is possible by understanding what customers want and need.

In addition there are also other application areas of customer analytics which do not match any specific dimension of CRM. For example Spangler et al. (2009, p. 243) write that advanced text analytics can be used to mine online data such as blogs, news forums, message boards, and web pages for customer insight. Companies can for instance analyze this online data in order to generate early warnings from brand and reputation issues. This is beneficial as the brand image and reputation are really important for consumer facing companies and because a brand can become easily negatively associated with an industry, environmental, or social issues. (Spangler et al. 2009, p. 243.) Other additional application areas of customer analytics include but are not limited to for example product and service development, sales forecasting, analyzing marketing campaigns and their success, and so on (Rygielski et al. 2002, p. 488; Arantola 2006, p. 79-87).

Customer lifetime value

Ngai et al. (2009, p. 2595) write that the customer lifetime value analysis is commonly defined as a prediction of the total net income that a company can expect from a specific customer. As already discussed shortly regarding customer profitability, there exist different models for analyzing customer lifetime value with data mining. Both customer profitability and customer lifetime value can be used for making decisions regarding the allocation of the company efforts (Wang & Hong 2006, p. 718). These include among other things marketing, sales, and service efforts. Stone et al. (2004, p. 125) and Wang & Hong (2006, p. 718) argue that highly profitable individuals or segments might need to be treated differently compared to other customers. The effort and costs related to unprofitable individuals or segments on the other hand should be reduced. In some cases unprofitable customers could even be encouraged to leave (Stone et al. 2004, p. 125). Arantola (2006, p. 87) states that differentiating customers for example based on their profitability can be used in sales, marketing, services, and product development.

One of the main ideas behind customer profitability and customer lifetime value is to identify those customers which are too costly to serve. Stone et al. (2004, p. 125) write that for example in business-to-business markets some large customers might be very demanding in terms of special terms and conditions or special prices. Similarly, in consumer markets some low revenue customers might be very demanding.

Studying the customer profitability further, Arantola (2006, p. 127) argues that customer profitability calculation is also more and more often used as a tool to improve company's profitability. The basic idea behind this is that a specific customer is not necessarily unprofitable but instead the company's approach towards that customer is not suitable. For example actual and potential customer profitability, when combined with the understanding of customer's needs, can be used for creating a segmentation model where correct products, prices, and contact channels are used with correct customer segments increasing the profitability of those customer segments in question (Arantola 2006, p. 88).

Product associations

Another important application area of customer analytics is related to transactions and buying behavior. One example of such applications is Market Basket Analysis (MBA). It is a modeling technique which is based upon the theory that certain groups of items are linked to other groups of items. For example a retail store customers might purchase milk and cheese together or bank customers might use specific bank services jointly (Chen et al. 2005, p. 339). In other words, if a customer buys a certain group of items, he is also likely to buy another group of items (Arantola 2006, p. 24).

The data used in MBA is collected from point of sales. Arantola (2006, p. 24) writes that all individual baskets or receipts from point of sales are analyzed and checked for product associations. Chen et al. (2005, p. 352) write that MBA which is also known as association rule mining is a useful method for extracting co-occurrences or associations from companies' transactional databases in order to discover customer purchasing patterns. Arantola (2006, p. 24) argues that MBA is especially interesting if the point of sale data can be collected from different stores of the same chain or from different chains of same company. Ngai et al. (2009, p. 2595) state that MBA is especially used for maximizing the customer transaction intensity and value. The insight regarding product associations can be used for example in designing websites, store layout, product mix and bundling, and other similar marketing strategies (Chen et al. 2005, p. 339).

When different product associations are known, different web stores and physical store layouts can be designed so that the revenue of a customer visit is increased. If two separate product groups sell well together, it might for example be a good idea to position those product groups next to each other. Similar design decisions can be made for web stores. When a customer buys a specific product, he can be immediately targeted or prompted with products which have been identified to sell well together with the product he just bought. On top of that, when the customer is identified, and therefore the segmentation is available, even more suitable products can be targeted or prompted based on the available transaction history and segmentation information.

The product associations can also be used for maximizing the impact of marketing activities and campaigns by examining which joint product offers are most likely to generate additional sales (Chen et al. 2005, p. 339). When the comprehensive customer data including segmentation information is combined with the product associations, the customers can be more effectively offered with other identified associated products or upgrades to the products they have already bought. This makes it possible to identify what kind of product combinations appeal to certain types of customers. Therefore it is possible to identify preferences of specific customer segments. This information can then be used to increase marketing effectiveness as the correct product offers or promotions can be targeted to correct types of customers. Arantola (2006, p. 70) argues that especially by analyzing behavioral data companies can predict with good accuracy what is the next product or a service to offer to a specific customer.

Customer needs and propensity to purchase

Data mining applications and tools allow companies to predict the probability for a specific customer buying a certain product (Turban et al. 2008, p. 136-137). For example neural networks can be used for predicting if customers would or would not buy a specific product based on the past behaviour of customers (Stone et al. 2004, p. 155). In other words, customer’s propensity to purchase a specific product or service is predicted. For example Kamakura et al. (2003, p. 46) state that with data mining it is possible to predict the best potential customers for cross-selling so that only those services, which the customer is very likely to be interested in, are offered. This can be done by analyzing the behavioural patterns of all customers.

Also external data gathered from a sample of customers with a survey can be used for enriching the analysis. These predictions can be used for creating targeted marketing including those products that a specific customer is likely to need and therefore buy. The same applies for example for web stores. Cho et al. (2002, p. 341) argue that by analyzing the web store transactions and click-stream data it is possible to learn the different product associations and customer preferences. These can then be used to give automatic product recommendations for customers based on their previous transactions and web behaviour. Also Fayyad (2003, p. 2) writes that by analyzing click-stream data, or in other words the navigational steps customers took to find what they desired, companies can understand customers’ decision making processes and use that insight to influence the future purchase decisions.

Another element which can be used when giving product recommendations and creating target marketing is sequential purchases. Customers often make sequential purchases by buying multiple products or services from the same company in a specific order so that the purchase of a certain product precedes the purchase of another product (Li et al.

2005, p. 233). For example Li et al. (2005, p. 234-235) propose a model which identifies these sequential purchases. This insight can then be used for example for identifying cross-selling opportunities (Li et al. 2005, p. 237-238).

Besides identifying the customer's propensity to purchase different products Stone et al. (2004, p. 125) write that it can also be relevant to understand customer's overall propensity to buy from the company. This is mainly due to the fact that it tells the "share of wallet" the company has achieved compared to its competitors.

Social network analysis

In addition to the other types of customer analyses, especially telecommunication companies have large possibilities to optimize their marketing and sales activities through social network analysis. This is possible because telecommunication companies have large quantities of data available from the usage of mobile phones. By analyzing this data these companies can identify the group leaders of social networks. Arantola (2006, p. 71) writes that these group leaders of social networks are important because it is believed that recommendation based marketing where the group leader of social network recommends a specific product or service to other members of that social network is multiple times more effective than even targeted marketing.

When social network analytics is combined with customer profitability information, it is possible to identify those social networks and their group leaders, who should be targeted with up-sell and cross-sell offers, in order to maximize the revenue impact. When social network analytics is combined with churn prediction, it is possible to identify those group leaders of social networks who have highest propensity to churn. As the loss of these customers could also lead to the loss of other members of those social networks, it is especially important to retain them.

4. BUSINESS INTELLIGENCE ARCHITECTURE MATURITY

This chapter introduces firstly three different business intelligence architecture related maturity models. The maturity stages and the different maturity categories of these models are also described. Especially the different business intelligence architecture related categories are focused on when the models are introduced. Secondly a new business intelligence architecture maturity model is introduced in this chapter. This new model is composed from the three existing maturity models in order to create a maturity model, which takes into consideration all the aspects of business intelligence architecture, including business analytics related technologies. As discussed in chapter 2, both the application of business intelligence and the application of business analytics require functional business intelligence architecture.

4.1. TDWI's business intelligence maturity model

In 2004 the Data Warehousing Institute (TDWI) created a business intelligence maturity model as an answer to the requests of business intelligence and data warehousing professionals. Since that the maturity model has been used by numerous companies in order to define their current business intelligence maturity. (Eckerson 2007a, p. 3). The maturity model consists of eight categories which are *scope*, *sponsorship*, *funding*, *value*, *architecture*, *data*, *development*, and *delivery*. The five maturity stages are *infant*, *child*, *teenager*, *adult*, and *sage*. This business intelligence maturity model is visualized in Figure 10. From the point of view of business intelligence architecture the categories architecture, scope, data, and delivery are especially interesting. Also value category is significant considering the research objective of this research.

In TDWI's business intelligence maturity model architecture category primary defines different maturity stages for company's data integration capabilities and data warehouse, or in other words repository, architecture. Scope category among other things further defines data warehouse architecture and the scope of business intelligence use. Data category instead includes maturity stages for data management related capabilities. The different maturity stages of delivery category define what kind of analytical and presentation applications and tools a company has and therefore what kind functionality and insight it is able to deliver. Value category and its maturity stages on the other hand define what kind of value the functionality and insight delivered can be used to achieve.

Category \ Stage	Infant	Child	Teenager	Adult	Sage
Scope	Individual	Department	Division	Enterprise	Inter-enterprise
Sponsorship	Non-existent or uncommitted	⟵⟶	Somewhat committed & accountable	⟵⟶	Very committed & accountable
Funding	None	Departmental budget	Divisional budget	Corporate IT budget	Self-funding
Value	Cost Center	Tactical	Mission critical	Strategic	Complete differentiator
Architecture	Spreadmarts	Non-integrated data marts	Non-integrated data warehouses	Central DW with or without data marts	BI or data service via service-oriented architecture
Data	Non trustworthy, not timely, not comprehensive	⟵⟶	Somewhat trustworthy, timely, and comprehensive	⟵⟶	Fully trustworthy, timely, and comprehensive
Development	Non-standardized processes	⟵⟶	Somewhat standardized processes	⟵⟶	Fully standardized processes
Delivery	View static reports	Analyze trends and issues	Monitor processes	Predict outcomes	Automate processes

Figure 10. TDWI's business Intelligence maturity model (adapted from Eckerson 2007a, p. 13).

Eckerson (2007a, p. 4) defines that an infant stage company mainly relies on operational reports produced by operational systems, instead of using specific reporting or analytical applications and tools. These operational systems usually store only that system's data and therefore the operational reports only show a limited set of data. Thus the operational reports often lack comprehensive and trustworthy data. (Eckerson 2007a, p. 13). When the data is not trustworthy also the obtained information might be misleading or inaccurate. Operational reports are also static and inflexible. If a new custom report is needed, it needs to be coded.

Furthermore, Eckerson (2007a, p. 4-5) writes that because of the inflexibility and non-trustworthiness of operational reports, infant stage companies also store their data in spreadmarts which are spreadsheets or desktop databases created by users who collect, clean, transform, aggregate, and format data themselves for individual or group consumption. The users, typically high-priced business analysts, who create spreadmarts and use them to fulfill their reporting needs basically perform all the functions of a data mart or data warehouse (Eckerson 2007a, p. 5-6). Data quality is another issue that companies in infant stage face. The data quality in operational systems or even in spreadmarts is poor or not well understood (Eckerson 2007a, p. 6). This results into an obstacle that needs to tackled with proper data management capabilities.

In child stage some of company's departments start to use data marts to help in different processes or to answer different problems instead of using huge amount of their business analysts' valuable time (Eckerson 2007a, p. 6). These data marts are usually departmental and they are not integrated or aligned with each other. Therefore the data is more comprehensive than in spreadmarts but it still is not nearly comprehensive enough. Often also data quality is poor and thus the data is not trustworthy. (Eckerson

2007a, p. 13). Eckerson (2007a, p. 6) states that in child stage companies start to use their first business intelligence technologies to create mainly ad hoc queries, standard reports, and to perform online analytical processing (OLAP). Especially using OLAP technologies allows these companies to increase their delivery maturity by analyzing different trends and issues. The main focus is to gain new insight by understanding past business situations better. (Eckerson 2007a, p. 6, 13) However, as stated, the departmental data marts basically have the same data management, data integration and data warehouse architecture problems as spreadmarts had. Inconsistent data definitions, data quality issues, and non-integrated data limits the insight acquired with the use of business intelligence technologies.

Eckerson (2007a, p. 7) states that in teenager stage the business units of companies have understood the importance of consolidating all the different departmental data marts into a single data warehouse. However, all the different business units still have their own data warehouses. This kind of data warehouse architecture provides a comprehensive view of business unit level data. This data is also becoming more and more trustworthy as multiple different data sources have been integrated into one creating more consistency (Eckerson 2007a, p. 13). The company is not still able to create a comprehensive and trustworthy view of data across the whole enterprise but it is able to create understanding and analyze its business better (Eckerson 2007a, p. 7). As the business unit level data integration capabilities have improved more timely information is also available. This makes it possible to monitor current state of business better than before through different alerts and dashboards (Eckerson 2007a, p. 7).

Investigating the model further, the companies in adult stage have been able to create unified data warehouse architecture (Eckerson 2007a, p. 8). This architecture includes a central data warehouse which contains a consolidated set of data from across the whole enterprise. All the data which company needs in order to report and analyze its business is located in the data warehouse. The data is also timely available with the help of highly capable data integration. Depending on the need, the data might be fed in real-time or in batches. At the same time, also the analytical applications and tools in use are more sophisticated. Adult stage companies use different analytical technologies and models to create forecasts, extrapolations, and predictions which can be used to anticipate needed business activity. (Eckerson 2007a, p. 9).

Finally, in sage stage the central data warehouse contains fully trustworthy, timely, and comprehensive data and the differences compared to adult stage are more related to the way company applies the analytical applications and tools in its business processes (Eckerson 2007a, p. 13). Through service oriented architecture (SOA) the company makes different kind of data and business intelligence services available to all applications. This makes it possible to used different analytical capabilities, such as rule engines and predictive models, in sophisticated composite applications in order to monitor and execute business processes in real time. In other words, sage stage companies are able to use trustworthy analytical capabilities to automate its business processes. (Eckerson 2007a, p. 10).

In these different stages of business intelligence maturity companies rely on business intelligence to answer different types of problems or questions. Therefore the insight acquired also varies through the different stages. In infant and child stages companies analyze historical trends in order to build awareness and understanding of the business. In teenager and adult stages companies exploit right-time information and analytical applications to work proactively to solve problems and optimize performance. In the sage stage companies use analytical applications with statistical models to automate decisions and processes. In the same way as insight also the business value of business intelligence changes through the different stages. During the infant and child stages the business value increases slowly as companies struggle to create a consistent view of data by consolidating departmental and individual silos and to provide proper business intelligence tools. The business value increases faster in teenager, adult, and sage stages as business intelligence starts to deliver a competitive advantage through rich insight and to drive mission-critical processes. (Eckerson 2007a, p. 4-13; TDWI 2009).

4.2. Davenport's and Harris' analytical capability maturity model

Davenport & Harris (2007, p. 114) have defined a five-stage maturity model around company's analytical capability. This maturity model is illustrated in Figure 11. The maturity stages from the lowest stage of maturity to the highest are *analytically impaired*, *localized analytics*, *analytical aspirations*, *analytical companies*, and *analytical competitors*. The maturity model considers analytical capability from *organization*, *human*, and *technology* perspectives. The maturity of organization's analytical objectives and processes, employee's skills, sponsorship, culture, and data and analytical technologies all affect the company's total analytical capability related maturity (Davenport & Harris 2007, p. 114). Especially the technology perspective with its data and analytical technologies related maturity stages is interesting considering analytics, business intelligence architecture and the focus of this research. Also the organization and especially its *analytical objective* perspective is relevant considering the research objective of this research. However, it is important to note that Davenport & Harris (2007, p. 128) state that analytical technologies alone are not sufficient to transform an organization. Instead, that requires all the different perspectives of analytical capability.

In Davenport's and Harris' analytical capability maturity model technology perspective defines widely the maturity of company's business intelligence architecture. It considers business intelligence and analytic technologies, repository architecture, data integration, and data management related topics such as data quality and data sourcing. The organization perspective considers among other things the use of data and analytical technologies and the insight obtained through analytical capabilities.

Stage		1 Analytically impaired	2 Localized analytics	3 Analytical aspirations	4 Analytical companies	5 Analytical competitors
Organization	Analytical objective	Limited insight into customers, markets, competitors.	Autonomous activity builds experience and confidence using analytics; creates new analytically based insigts.	Coordinated; establish enterprise performance metrics, build analytically based insights.	Change program to develop integrated analytically processes and applications and build analytical capabilities.	Deep strategic insights, continuous renewal and improvement.
	Analytical process	Doesn't exist.	Disconnected, very narrow focus.	Mostly separate analytic processes. Building enterprise-level plan.	Some embedded analytics processes.	Fully embedded and much more highly integrated.
Human	Skills	None.	Pockets of isolated analysts (may be in finance, SCM, or marketing / CRM).	Analaysts in multiple areas of business but with limited interaction.	Skills exist, but often not aligned to right level/right role.	Highly skilled, leveraged, mobilized, centralized, outsourced grunt work.
	Sponsorship	None.	Functional and tactical.	Executive – early stages of awareness of competitive possibilities.	Broad C-suite support.	CEO passion. Broad-based management commitment.
	Culture	Knowledge allergic – pride on gut based decisions.	Desire for more objective data, successes from point use of analytics start to get attention.	Executive support for fact-based culture – may meet considerable resistance.	Change management to build a fact-based culture.	Broadly supported fact-based culture, testing and learning culture.
Technology	Quality data & analytical technologies	Missing / poor-quality data, multiple defines. Unintegrated systems.	Recent transaction data unintegrated, missing important information. Isolated BI / analytic efforts.	Proliferation of BI tools. Data marts / data warehouse established / expands.	High-quality data. Have an enterprise BI plan / strategy, IT processes, and qovernance principles in place.	Enterprise-wide BI / BA architecture largely implemented.

Figure 11. Davenport's and Harris' analytical capability maturity model (adapted from Davenport & Harris 2007, p. 114).

Firstly, Davenport & Harris (2007, p. 156) state that companies which are in analytically impaired stage have not integrated their systems and they use poor-quality data with multiple different defines. In many cases the needed data is even missing which calls out for gut based decisions (Davenport & Harris 2007, p. 114). These companies need to focus on improving their transactional systems in order to have the consistency and quality of their data up. In other words these companies lack the prerequisites which are needed for analytics and therefore need to focus first on these prerequisites. (Davenport & Harris 2007, p. 108). If these companies have any analytical initiatives, they are tactical in nature and have only limited impact on business. From the technology perspective this is resulted by the missing or poor-quality data which can only be used to produce limited insights into customers, competitors, and markets.

Secondly, in localized analytics stage companies collect transactional data quite efficiently into their transactional systems, but especially the recent transactions are not integrated between different systems. Often also the right data for better decision making is still missing and the starting business intelligence and analytic efforts are isolated into certain departments. (Davenport & Harris 2007, p. 114, 156). These departments are however usually able to implement small stand-alone analytical tools and prepare the departmental data into such a shape that they can produce new analytically based insight. This insight can be used to achieve measurable benefits such as productivity gains or cost savings. (Davenport & Harris 2007, p. 117-118). Davenport & Harris (2007, p. 121) note that one of the risks in this second stage is that the departments with business intelligence and analytical efforts have started to use different hard-to-integrate tools and data sets which will cause problems in future.

Furthermore, Davenport & Harris (2007, p. 114) mention that companies which are in analytical aspirations stage have further acquired business intelligence technologies and established their first data warehouse, data marts or both. These companies have launched their first major initiatives to use sophisticated analytics in addressing strategic business problems (Davenport & Harris 2007, p. 121). However most of the data these companies possess still is not integrated, accessible or standardized (Davenport & Harris 2007, p. 156). This results to the fact that for example a complete picture of a customer cannot be acquired as different lines of business still have their own and not shared customer data.

Contrarily, Davenport & Harris (2007, p. 114, 156) state that in analytical companies stage companies have mainly been able to fix their data related issues and can now use high-quality data as a basis of their analytical activities. The high-quality data is created by integrating and standardizing data across the company. In these companies sophisticated analytical technologies are used on enterprise level instead on departmental or business unit level. The world-class analytic capabilities provide large amount of strategic insight and new ideas which can be used for competitive advantage. (Davenport & Harris 2007, p. 124).

Lastly, the analytical competitor stage companies have largely implemented an enterprise-wide and highly sophisticated analytic architecture which is fully automated and integrated into processes (Davenport & Harris 2007, p. 114, 156). These companies have been able to define what data they need for analytics, how much they need it, and where that data is located (Davenport & Harris 2007, p. 159). They have also been able to acquire high-quality data which is correct, complete, current for the business problem in hand, consistent throughout the organization, enriched with metadata, and controlled (Davenport & Harris 2007, p. 163). For these companies the analytical capabilities are the key to their strategy and competitive advantage instead of just being very important (Davenport & Harris 2007, p. 125).

4.3. LaValle's business analytics and optimization maturity model

LaValle has taken a different approach related to the business intelligence and business analytics maturity. In his study report LaValle (2009, p. 2-3) presents a maturity model related to business analytics and optimization. This includes both how a company manages information and how it applies that information. The categorization was created based on a survey where nearly 400 business leaders worldwide were asked about their use of information and the application of business intelligence. This categorization is illustrated in Figure 12.

LaValle's maturity model consists of two axes and five maturity stages. The axes measure the maturity of *information and analytics* and *business operations* dimensions. Information and analytics dimension considers how the business manages information and learns from it (LaValle 2009, p. 2-3). Business operations dimension on the other hand considers how the business applies information to achieve its goals (LaValle 2009, p. 2-3). The maturity stages from the lowest stage of maturity to the highest are *ad hoc*, *foundational*, *competitive*, *differentiating*, and *breakaway.* In his study LaValle mainly focuses identifying the differences between the characteristics of top performers[1], or so called breakaways, and the lower performing organizations instead of defining the characteristics of organizations in each of the maturity levels.

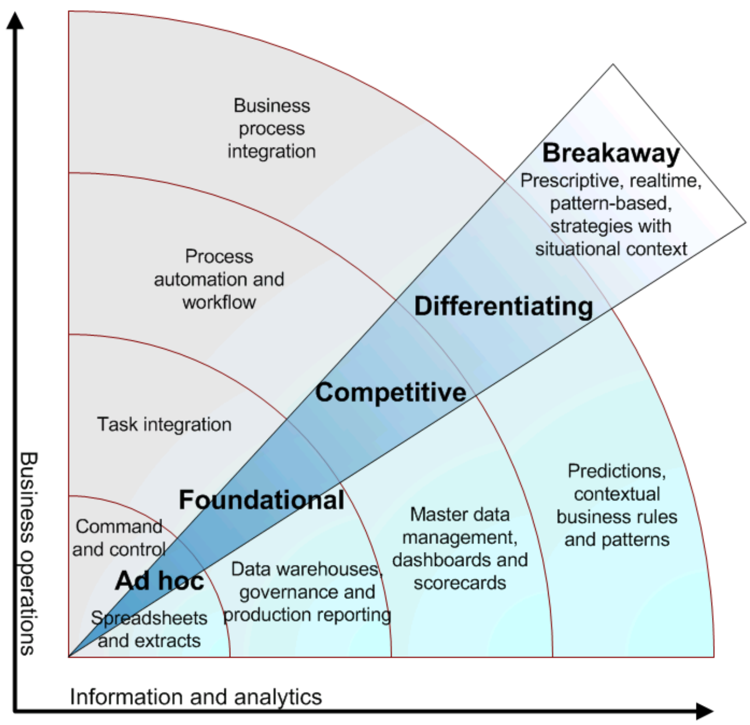

Figure 12. LaValle's business analytics and optimization maturity model (adapted from LaValle 2009, p. 3).

As stated LaValle looks business analytics and optimization maturity from the perspective of two different dimensions. When these dimensions are studied more closely, business operations dimension includes subjects related to policies, business processes, and organization. Therefore this dimension considers many same subjects as

---

[1] The respondents in LaValle's study classified themselves compared to their industry peers. Based on the results the respondents were put in two groups: top performers (upper 20 percent) and lower performers (lower 40 percent).

the organization and human perspectives of Davenport's and Harris' analytical capability maturity model. Information and analytics dimension instead includes many same things as the technology perspective of Davenport's and Harris' analytical capability maturity model. Thus, as in the technology perspective, many of the subjects included in information and analytics dimension are closely related to the different aspects of business intelligence architecture maturity. As that is the main focus of this study, when considering the maturity models, it is beneficial to understand the related characteristics LaValle's study reveals about the top performers. At the same time it is important to also understand that if insight is not applied correctly, its potential is not fully harnessed. That is where the business operations dimension comes into the picture. LaValle (2009, p. 4) states that in order to create new intelligence and insight business analytics and optimization top performers use both state of the art toolsets and processes to understand, share and analyze information.

Investigating LaValle's study further, it revealed that twice as many top performers were in their own opinion aware of the full scope of information available in their operations and how they can use it precisely. These organizations actively collect, monitor, and use information from inside and outside of the organization. (LaValle 2009, p. 4.) The top performers also used three times more often sophisticated data governance systems when compared to the lower performers. With data governance LaValle considers among other things processes, master data management, and common data definitions. The most sophisticated organizations possessed strong processes and management systems with automated data governance tools. (LaValle 2009, p. 5.)

LaValle's study evaluated also the differences in the usage of various business analytics and optimization toolsets between top and lower performers. These toolsets included dashboards and visualization, analytical and predictive tools, business rules management, content management, data integration, and master data management (LaValle 2009, p. 6). As an example, top performers rated their predictive dashboards and visualization tools as world-class tools over four times more often than lower performers. Same applies for analytic and predictive tools and for business rules management tools which both were rated as world-class tools by top performers over two times more often than lower performers. Similar differences were noticed also regarding tools that are used to acquire trusted information. Tools which are used among other things to distribute content, to assure the quality and usability of information, and to provide a single view of the truth were rated more favorably over two times more often by the top performers. These tools include content management, data integration, and master data management tools. (LaValle 2009, p. 7.)

"In short, effective tools – from content management to dashboards to visualization – go a long way in helping to become a breakaway organization. They allow the entire organization to anticipate and challenge – while at the same time providing safeguards for effective management systems." (LaValle 2009, p. 7.)

4.4. Combined business intelligence architecture maturity model

The concept of business intelligence architecture was defined in section 2.3. Furthermore, as the maturity models which were introduced in the previous sections did not cover all the different aspects related to business intelligence architecture this research uses a composite maturity model which is illustrated in a high level in Figure 13. In additions, a more detailed maturity model including all the characteristics of companies in different stages of business intelligence architecture maturity is described in appendix 1.

The maturity model in question is composed from all the three previously presented maturity models. These maturity models are TDWI's business intelligence maturity model, Davenport's and Harris' analytical capability maturity model, and LaValle's business analytics and optimization maturity model. All the three maturity models support each other without having conflicts between each other. As they all look at the maturity from a little bit different perspectives they therefore supplement each other.

Analyzing the three models further, the TDWI's maturity model considers many of the different aspects of business intelligence maturity such as data management, data warehousing, and business intelligence tools and applications. However, it lacks a deeper consideration of analytical tools and applications, which would be helpful as this research is focused especially on customer analytics.

Secondly, Davenport's and Harris' analytical capability maturity model concentrates on aspects such as business intelligence and analytical technologies, repository architecture, data integration, and data management. Compared to TDWI's maturity model especially the aspects related to analytical technologies and some aspects of the data management such as data sourcing and quality add value.

As a third complementary maturity model, LaValle's study and business analytics and optimization maturity model give concrete examples from the differences between higher and lower maturity stages. LaValle's model studies many of the same aspects as Davenport's and Harris' maturity model with an especially keen focus to the technologies and tools involved in different parts of business intelligence architecture.

Next the categories and stages of the business intelligence architecture maturity model are introduced in more detail. This is beneficial as the business intelligence architecture maturity model is used to evaluate the maturities of different case companies later on in this study.

Category \ Stage	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(A) Data sourcing	Not comprehensive	↔	Somewhat comprehensive	↔	Fully comprehensive
(B) Data quality & integration	Non-trustworthy and not timely data	↔	Somewhat trustworthy and timely data	↔	Fully trustworthy and timely data
(C) Data warehouse architecture & scope	Individual spreadmarts	Deparmental non-integrated data marts	Business unit/division level non-integrated data warehouses	Central enterprise data warehouse with or without data marts	Data service via service-oriented architecture with inter-enterprise scope
(D) Analytical technologies	None	Isolated analytic efforts and stand-alone analytical tools	First initiatives to use sophisticated analytics	Sophisticated analytical technologies	Highly sophisticated analytical technologies
(E) Business intelligence technologies	Only operational reports and extracts	Isolated BI efforts and stand-alone reporting tools	Proliferation of BI tools	Centralized BI tools	BI / BA via service-oriented architecture

Figure 13. Combined business intelligence architecture maturity model (partly adapted from Eckerson 2007a; Davenport & Harris 2007; LaValle 2009).

An *infant* stage company uses (Category A, Data sourcing) only individual operational systems with limited set of data as internal data sources. (Category B, Data quality & integration) These systems are not integrated to other systems and they contain non-trustworthy data. Because of data quality issues and inflexibility of operational systems (Category C, Data warehouse architecture & scope) individual users create spreadmarts to fulfill their reporting needs. These companies (Category D, Analytical technologies) do not have any analytical technologies and they (Category E, Business intelligence technologies) rely on spreadsheet based reporting or on the operational reports and extracts produced by operational systems.

When a company has reached the *child* stage, (A) some of the company's departments use few integrated operational systems as internal data sources. Data from these systems is integrated to (C) departmental data marts, which usually are not integrated to other data marts. (B) This integration is done occasionally with company's first ETL-processes providing also slight data quality improvements. In addition, child stage companies have possibly (D) started to use first stand-alone analytical tools with functionality such as OLAP. On top of this, they are (E) using their first business intelligence technologies allowing them to create standard reports and ad hoc queries.

Furthermore, a *teenager* stage company has (C) business unit level non-integrated data warehouses and more consistent data definitions compared to those of companies in previous stages. (A) These data warehouses hold comprehensive business unit level data combined from multiple internal data sources. (B) With improved data integration and master data management capabilities the data is integrated in a more timely fashion into the data warehouse and it is more trustworthy. These companies also (D) have started to use their first sophisticated analytical technologies such as data mining tools, which are able to identify patterns from data. At the same time (E) different business intelligence tools with additional functionalities such as dashboards and scorecards have proliferated.

Those companies which are in the *adult* stage have been able to create (C) unified and centralized data warehouse architecture used throughout the enterprise. The data warehouse contains (A) consolidated enterprise level data from large amount of internal and also some external data sources, such as market researches. (B) Data integration is done with highly capable tools which allow both real-time and batch based integration depending on the need. Also issues related to data quality are mainly fixed. In addition, adult stage companies use (D) sophisticated analytical technologies which are able to make different predictions widely. Furthermore, (E) the business intelligence tools are centralized and used for instance through common portal throughout the enterprise.

Companies which have been able to reach the *sage* stage are (A) aware of the full scope of available information and use fully comprehensive data from large amount of integrated internal and external data sources. This data is (B) fully trustworthy and timely as these companies have sophisticated ETL-processes and master data management. The unified and centralized data warehouse architecture has (C) an inter-enterprise scope providing data services for any applications. Also the data definitions are consistent throughout the organization. These companies are also using (D) highly sophisticated analytical technologies including functionalities such as predictive dashboards and business rules management, which can be integrated as part of business processes. The used (E) business intelligence technologies make it possible to provide business intelligence services for any applications and also to partners and customers.

As do all the three maturity models from which the business intelligence architecture maturity model was composed, so does this maturity model suggest that the business value, enabled by the companies' business intelligence architectures, is increased on higher maturity stages. Comparing the stages from the business value point of view, infant stage companies are lacking many of the needed prerequisites and are therefore able to only create limited insight into customers, competitors, and markets. This insight can be used to impact business only a little. Child stage companies are able to understand the past business situations better by analyzing historical trends and issues. This new insight allows the companies to achieve some measurable benefits such as productivity gains or cost savings. Teenager stage companies can already address strategic business problems with proactive responses by monitoring the current state of the business better than before. They are also able to create some strategic insight by identifying unknown patterns from the data. Adult stage companies are able to create strategic and automated operational insight by identifying patterns and making predictions. This can be used to achieve strategic value. Lastly, sage stage companies have reached a level where analytics is used to monitor, optimize, and automate decisions and business processes such as customer interactions. This allows sage stage companies to achieve competitive advantage.

5. RESEARCH METHOD AND MATERIAL

In this chapter the research method used in this research is discussed through. Secondly, the data collection method and research material are described. Thirdly, the data analysis methods are presented in detail.

5.1. Research methods

There are multiple different research strategies which can be used while doing a research. Some of these research strategies include for example experiment, survey, archival analysis, history, and case study (Yin 2003, p. 5). Additionally, each different research strategy is suitable for multiple situations and in certain type of researches the researcher has multiple options to pick from. There however are good guidelines which help researchers to identify a suitable research strategy. Yin (2003, p. 5) defines that there exists three conditions for research strategy selection. These conditions are the type of research question, the degree of control the researcher has over the behavioral events related to the research, and the focus on contemporary versus historical events. These conditions are shown in Table 1. When research strategy is selected one should evaluate all these three different conditions.

Table 1. Guidelines for choosing research strategy (adapted from Yin 2003, p. 5).

Strategy	Form of Research Question	Requires Control of Behavioural Events?	Focuses on Contemporary Events?
Experiment	how, why	yes	yes
Survey	who, what, where, how many, how much	no	yes
Archival analysis	who, what, where, how many, how much	no	yes/no
History	how, why	no	no
Case study	how, why	no	yes

Yin (2003, p. 5) states that research questions can be categorized into five different categories. These categories include “who”, “what”, “where”, “how”, and “why”. Each research question category matches certain types of research. Furthermore, a research can be descriptive, explanatory, or exploratory in nature (Yin 2003, p. 6-7).

In this research the research questions are either “what” or “how” questions. However, as already stated in chapter 1 the objective of “what” and some of the “how” questions are to get a preliminary understanding of the research area and to build a framework for conducting the empirical research. This understanding is acquired in the theoretical part of this research. Therefore the research question which is answered in the empirical part of the research is a “how” question. This question is:

How does the maturity of different aspects of business intelligence architecture affect the insightfulness of customer analytics?

Considering the type of the research question three different research strategies are possible. These are experiment, history, and case study. (Yin 2003, p. 13.) In this research the focus is however on the current events; that is the current maturity of company's business intelligence architecture and the current insights obtained through customer analytics. On top of that, when the researcher's possibility to manipulate the relevant behaviors and the phenomenon studied is evaluated, it can be stated that the researcher has very little or no possibility to do that. Therefore the case study research strategy is appropriate for this research. Yin (2003, p. 13) writes that case study is an empirical inquiry where a specific contemporary phenomenon is studied in its real-life context. Furthermore, Yin states that case study is especially useful if the boundaries between the phenomenon and context are not clearly evident. In addition, Yin (2003, p. 15) argues that case study is especially good in explaining presumed causal links in real-life interventions. Thus, case study is a suitable research strategy for a research where the main idea is to explain the causal link between the maturity of business intelligence architecture and the insightfulness of customer analytics.

Multiple case companies were studied in this study. This was done because the goal was to get a wider understanding about the researched phenomenon than could be acquired by studying only one case company. Therefore the actual research strategy of this research is a multiple-case study. Yin (2003, p. 46) writes that multiple cases can be used for example because the evidence from multiple cases can be considered more compelling. However, all the cases need to be selected carefully so that they either predict similar results or predict contrasting results but for predictable reasons (Yin 2003, p. 47). In this research the cases are predicted to be different but the reasons for this are predicted to be related to the business intelligence architecture maturity.

5.2. Research material and interviewees

Multiple different sources of evidence and data collection methods can be used in a case study. These sources of evidence include documents, archival records, interviews, direct observation, participant-observation, and physical artifacts. (Yin 2003, p. 13.) Yin (2003, p. 89) also states that interviews are one of the most important sources of information in case studies. Depending on the questions the interviewer prepares the interview can be open, semi-structured, or structured (Järvinen & Järvinen 2004, p. 145). In this research the main research material is collected through semi-structured interviews. Järvinen & Järvinen (2004, p. 145) state that both open themes and structured questions are included in a semi-structured interview. This is also the case with this research as two main themes, business intelligence architecture and customer analytics, were included into the interview. Both themes had few predefined high level questions which gave the basis for the interview. All in all, a semi-structured interview was selected as the data collection method because it was evaluated to be a good way to gather information about the phenomenon. This is especially true because the insightfulness of customer analytics is not something that can be easily quantitatively measured. Therefore the qualitative information gathered in the semi-structured interviews is a good way to create a picture about the phenomenon.

The research material includes five semi-structured interviews which were all recorded and conducted between 25.01.2011 and 07.03.2011. The companies and interviewees were selected based on the following criteria. First of all, all the companies are currently practicing business intelligence and there also exists at least preliminary activities related to customer analytics in each of the case companies. Depending from the case the interviewees were selected based on their position in the company. If possible, a person who had a wide view to both information management and business was selected. A good example of such a person is the head of business intelligence or customer analytics related function. If this was not possible, more than one person was interviewed from a single company so that sufficient views to both information management and business were acquired. Other criteria such as the size of the company, industry sector, age of the interviewee or gender of the interviewee were not used as they were assessed to be irrelevant for the research. The participation of companies from different industries was seen as good thing because it allowed the researcher to get a wider picture from the research topic. The title and the experience in business intelligence and business analytics of the interviewees and the case company's industry sector are listed in Table 2.

Table 2. Case company and interviewee details.

Case	**Industry sector**	**Interviewee's title**	**Interviewee's experience in BI & BA**
1	automotive	Business Intelligence Manager	about 5 years
2	retail	Customer Information Director	10 years
3	telecommunication	Consumer Analytics Director	over 5 years
4	retail	Information Services Director	over 10 years
		eCommerce Development Manager	over 3 years
5	telecommunication	Development Manager - Analytical CRM and Data Mining	over 10 years
		Senior Development Manager	over 10 years

All the interviewee's were given a base of the interview both in Finnish and in English beforehand. This base of the interview can be found from appendix 2.

5.3. The data analysis and analytical techniques

Yin (2003, p. 115) states that having a general analytical strategy is a best preparation for conducting case study analysis. One of these general analytical strategies is to create case descriptions which describe the cases. This research uses case descriptions and they can be found from chapter 6. All the case descriptions were written based on to the recorded semi-structured interviews. The records were listened multiple times through so that the case descriptions would describe the current business intelligence architecture and the used application areas of customer analytics reliably. These case descriptions were then analyzed based on to the preliminary understanding of the research area and the framework which was created in the theoretical part of this research for conducting the empirical research.

The main analytical technique used in this research is a cross-case synthesis. Yin (2003, p. 134) states that in a cross-case synthesis the different cases are compared with each other in order to find similarities. Possible methods for this are example tables where data from all the individual cases is displayed according to some uniform framework. In this research such comparison is first done for the business intelligence architecture maturities of the case companies. For this reason the maturity of each case company is first evaluated based on to the framework created in the theoretical part of this research. Secondly, similar comparison is done to the different application areas of customer analytics practiced in the case companies. And lastly, similar comparison is done to the level of customer analytics' sophistication between the case companies. Especially the correlation between the business intelligence architecture maturity, the amount of application areas of customer analytics, and the level of customer analytics' sophistication is studied by using these comparisons between case companies. This is done by analyzing the patterns found from each of the comparisons. Patterns in this context mean similarities in the changes of business intelligence architecture maturities between the case companies, amounts of application areas of customer analytics between the case companies, and the level of customer analytics' sophistication between the case companies.

6. CASE DESCRIPTIONS

In this chapter all the case companies are described based on the conducted semi-structured interviews. All the case descriptions follow the same formula. First both the case company and the interviewee are introduced and their experience regarding business intelligence and analytics is described. Secondly, the business intelligence architecture of the case company is discussed. This includes among other things data sourcing, data quality and integration, data warehouse architecture and scope, analytical technologies, and business intelligence technologies. Thirdly, and lastly, the different application areas of customer analytics practiced and the customer insight which the case company is able to create are described. Also the effect this customer insight has on to the company’s ability to increase and sustain sales is discussed.

6.1. Case 1

The first case is described based on to the interview of Business Intelligence Manager from a large automotive company. The interviewee works in the company's information technology department and leads a business intelligence team which includes a data warehouse manager and different persons from the business. The interviewee's tasks consist from the creation of business intelligence services for the business and from the development of business intelligence in the company. In the current position the interviewee started in October 2009 but has in addition worked in other positions in the case company already from year 2005. During this time the interviewee has worked in projects related to sales reporting portal and financial management systems. In addition, the interviewee has also worked as a controller and has therefore acquired a good understanding about the company's core business. Furthermore, the interviewee has experience from reporting, budgeting, forecasting, and analyzing tools.

The company defines business intelligence as reporting and analyze of business information. This also includes planning, budgeting and forecasting. Furthermore, also data warehousing is seen as a part of business intelligence. The company has worked with data warehousing and reporting for a quite some time already as the first data warehousing and reporting project was in 1997.

6.1.1. *Business intelligence architecture*

Currently the company has one centralized data warehouse, which is connected to multiple different source systems. These source systems include systems such as CRM system for business customers and partners, production system, warehouse system, and financial systems. In most cases data from these source systems is integrated with overnight ETL-batches into the centralized data warehouse. However, also the criticality of the data is considered in data integrations and therefore some of the data can also be loaded during days. Currently sales data is identified as the most critical data and for this reason it is kept topical also in the data warehouse. Furthermore, real-time data integrations have been considered but the current cycle of decision making does not really require real-time data at the moment. Instead of being a critical feature, real-time data integration is seen as a nice to have feature.

The company collects transactional information from its business customers and partners who sell the company's products to consumers. Consumers however are not identified and therefore the transactions cannot be connected to a specific consumer. Business customers and partners are however a different matter. The company has a lot of information about their business customers and partners. They for instance know which products were sold to a specific business customer in what particular time. In addition, they also store detailed contact history and by using this information the company has for example created detailed business customer segments. Additionally, also different documents related to the customers such as contracts are stored in CRM system but they are not connected into the centralized data warehouse. All in all, the company does not currently use unstructured data sources.

Part of the company also has a web store where they sell different products and services to consumers. This web store also includes a loyalty program for which the customers input their details into the system. However, even though the transactional data itself is collected to the data warehouse, the customer data is not at the moment used in the data warehouse. This is however something the company is considering fixing in the future so that it could understand the consumers better. Furthermore, the company does not currently collect or mine data from the unstructured content of internet. The marketing department follows the development of the company's brand but these activities are not at the moment anyhow related to the business intelligence environment. Instead, the company's brand is followed through market researches.

The quality of data in the data warehouse is followed with a manual process with specific monitored points. The general assumption is that the data quality should be good already in the source systems. Additionally, also the normal ETL-process and master data management handle some quality related matters. There however is not a specific tool for data quality management and currently the unfortunate fact is that some poor quality data streams trough to the end user interface.

Furthermore, when reporting is concerned, many capable tools have been given all the way to end users. OLAP based reporting and analysis has been practiced for a long time in the company which has caused the fact that there has never really been a standard reporting culture in place. Additionally, every business unit has started to practice reporting and analysis from their own perspective. Only OLAP cubes are done centrally while the actual reporting is mainly done by the power users or capable end users. All in all, there are many power users and capable end users in the company and the business intelligence related operating model is self-service business intelligence. In the interviewees opinion this has however been mainly possible thing because the current reporting tools are quite simple and easy to use. When new more advanced reporting and analytical tools are taken into use this will not be so easy any more and therefore at that time some kind of centralization needs to be practiced also in reporting and analysis side. On top of the different capable tools which have been given to the end users there exists also a reporting portal which is somewhat used. However, as said, also other reporting tools are used. Additionally, the reporting portal is not currently used for instance to offer reporting as a service to customers or suppliers.

The company does not currently use analytical technologies such as data mining tools. Therefore the main analysis is done with OLAP tools or with straight queries into the data warehouse. All in all, at the moment the company's focus is on enhancing the basic business intelligence environment so that it matches the current business and its strategies. After this foundation is ready the goal is to implement new tools and additional elements, which can be used to create new real business value, on top of it.

6.1.2. *Customer insight and sales*

The company uses multiple different factors when they create customer segments. These factors include also other data than only demographical data. Additionally, the company also tries to include the degree of loyalty into the segmentation. On top of this, the company also calculates business customer profitability and they are aware of the long term profit in the whole supply chain. The calculation of customer profitability is considered as an important thing and it is used in both sales and marketing. The profitability related information is used to for example direct sales to those customers who are most profitable. Currently the company is however in such a situation that all the production is bought and therefore the demand is higher than supply. Therefore the profit can naturally be maximized by targeting sales to the most profitable customers. Additionally, these most profitable key customers receive best service.

The company does not use analytical applications to predict when and what products a certain business customer needs as this prediction is instead done by the business and is considered as one of the core competencies of the business. The business in general also follows some cycles which can be used to help in this prediction. Sales as a whole is however forecasted regularly and the company has specific tools for this. Sales forecasting can be based on for example to the transaction history, discussions with the customers, and on the market understanding. The interviewee however states that it is important to notice that there are many manual phases included to this forecasting. Nevertheless, these forecasts are for instance used to focus sales related activities.

Furthermore, as a result of the operating model of the company there is no need to identify for example cross-sell opportunities for partners as each partner has written an agreement with certain product portfolio. All in all, it can be stated that using analytical applications and tools to really understand customer behavior is something that is still mostly in the future. Nevertheless, the reporting and OLAP has however reached a state where the company's information technology department does not have to justify the costs which arise from business intelligence. The value which is achieved with the help of reporting and OLAP is not really measured as it is considered to be really hard but instead the company as a whole has understood the value of information. Especially after the recession the value of critical information has become clear to all.

In interviewee's opinion the next thing which should be improved regarding customer analytics is definitely the understanding of customers because in the end they are the ones who use the company's products. On top of this, the interviewee also emphasized the understanding of future market potential and the understanding of customer's buying behavior. It would be beneficial to know what factors affect to the customers decisions and also how the customers see the company's brand. The interviewee sees that for example social media and the information which is available in there would be important in order to understand these things. Furthermore, the interviewee sees this as an especially important thing because nowadays the market is shaping the structures instead of companies. Companies cannot anymore just push what they produce into the markets.

6.2. Case 2

The second case is described based on to the interview of Customer Information Director from a large retail company. The interviewee and the interviewee's department work with different subjects regarding marketing and customer information. The department includes teams for customer information management, analytics, reporting, and market research. In addition, they also help other functions or processes with their analytical capabilities.

Currently the interviewee's main task is to change management processes as information based in the whole company. The interviewee is basically an internal management consultant, process developer, and an analyst who defines high level directions for analytics and also questions the current solutions. The goal of this is to link the company's strategies and operations tightly together. The interviewee has been in the current position from year 2001. On top of this, the interviewee has a vast experience from the industry as the interviewee has worked in many different positions in the industry's value chain. Furthermore, the interviewee has also used different analytical tools because the interviewee needs to understand what actually can be done with these tools.

At the time the interviewee started to work in the company customer information was not yet really used in any way in management. In 2003 the company took its first steps to acquire tools and better information which could be used to help in customer analytics.

6.2.1. *Business intelligence architecture*

The company has a few different data warehouses which store data for decision making. The company is however able to combine customer data, transactional data, market research, and other relevant sources of data so that a larger picture from customers can be created. The different data sources used include among other things CRM system and transaction systems. Additionally, the company also has a loyalty program which produces important customer related data. Also different external sources are used. One of the most important sources of external data is market researches which are for instance used to get comprehensive data from the actions of competitors. Furthermore, the external data is connected to the internal data because otherwise there would be a risk for circular arguments in decision making. For example also competitors' actions affect the behavior of customers and therefore all the changes in customer behavior might not only be related to the company's actions.

Furthermore, the data used in analytics and reporting is seen as relevant and this is ensured by closely considering all the data gathered. In other words only data which in valuable for business's value chain or for customers is gathered. All non-relevant or outdated data is put away. Data quality on the other hand is all the time improved and the data quality related processes are currently under automation. The company however sees that, even though not nearly all is ready yet, the data quality is at least sufficient and that there are no major problems. Furthermore, data integration to the data warehouses is currently done once a day with batches. In near future this will however most likely change so that data is available for reporting and analysis in more real-time.

Data integration, data modeling, and different semantic layers used in reporting and analytics are all managed centrally. The same centralized team also maintains all the analytical and reporting tools. In addition, the different business processes and departments have also employees who are responsible integrating the insight provided by the centralized analytical team into the everyday processes. All in all, the company has acquired much internal knowhow and internally managed tools as in its opinion the data modeling and other related activities require a deep industry understanding.

Furthermore, there is currently one main centralized analytical tool in use. This tool is connected to the other parts of the business intelligence environment and can also be integrated with other systems. In addition, also other analytical tools are used now and then if needed. Especially data mining is extensively practiced and the company actually sees all the different analytical techniques or methodologies, such as predictive analytics, simulation, rule engines, as data mining. The need in hand states the required technique and data. Text mining is however something that has not been used in the company yet but for example the unstructured data in the internet is seen as an interesting source of customer data.

The company uses also reporting tools for information delivery and has a centralized reporting platform. However, more and more often the reports are not enough. The extra information and insight needs to be integrated to the business processes in exactly correct format. This way it supports the decision making in optimal way. When such integration is done the persons executing business processes do not often even know that they are dealing with enriched information. Instead they are just using their everyday operational tools which have customer insight embedded. However, even though a lot of customer insight is embedded in some of the business processes the company identifies that there is still a lot to do.

6.2.2. *Customer insight and sales*

The interviewee states that customer information the company possesses is used to create customer insight and direct answers to relevant business questions. In addition, these answers are also brought straight to the management systems and processes if possible. This supports the company's target to be the best company in agile operations guidance and tailoring among its industry. Furthermore, the goal of customer analytics in the company is to improve the customer experience and competitiveness of the company though information. All things, which are done as a part of customer analytics, should help the business; not to create for example reports which cannot be easily used to make better decisions. The interviewee also emphasized that for example analytical tools as such do not create any customer insight. Instead they are tools which help in the creation of customer insight while the analytical processes are actually creating the insight.

What comes to the actual application areas of customer analytics the company creates for example different customer segments based on the backgrounds, consuming styles, values, events, and opinions of customers with analytics. The company also profiles customers in different areas. For example the customers in the neighborhood of a specific store can be profiled and that information can be used in the decision making of that store. In addition the company is also able to profile those customers which are competitors' customers.

The company also understands the different factors affecting customer loyalty and that these factors vary a lot between different types of customers. This understanding can be given as an input to the different processes which can be applied to increase customer loyalty. Additionally, the interviewee stated that the company uses for example recency, frequency, and monetary values when estimating the degree of customer's loyalty.

Integrating this insight as part of the operational processes is something the company sees as a next step regarding customer loyalty. In addition, also understanding of customer lifetime value is considered to be an important part of the bigger picture.

Furthermore, the company also identifies sales opportunities based on the customer segments and buying behavior of customers for example in certain area. They also try to identify trends which indicate that the demand of a certain product or product group is growing. When the offering of a specific store is decided, all this information plus the profile of the customers in that area, is used to help business in making good offering related decisions. What comes to targeted marketing, the company identifies those customer segments which should be targeted for a reason or another. The problem until today has been to identify what actually to include in to the message delivered to those customers. Recently improvements also in this have been achieved and the marketing is more genuinely targeted based on the needs of that customer segment. The next step is to bring the targeting fully to individual level.

In the interviewee’s opinion especially any kind of changes in the life situation of a customer can be used to indentify different possibilities. For this reason the company already sends individually targeted marketing when certain customer related events happen. For example changes in customer address can be used to trigger automated and targeted marketing. The interviewee however states that it is important to act at the right time. If the marketing activity occurs too quickly or too slowly after a specific customer event, the effects of the marketing can be decreased. All in all, the company is more and more taking different triggered marketing actions in use. These can include for example marketing based on predicted customer churn, identified up-sell possibility, and so on.

On top of practicing customer analytics the company also uses their analytical capabilities among other things for fraud detection, and supply chain optimization. And as stated, all the insight, which is created by the different analyses, is connected into business processes if possible. Because of this, among other things, the interviewee’s department has acquired a degree of trust where they do not have to anymore justify their actions to management. Instead they can just focus on improving their services and doing the things they have planned.

In the future the company will focus at least on to the possibilities related to web mining, which enable the analysis of the unstructured data in the internet. In addition the interviewee states that also the possibilities to provide tailored information and services to the customers through different mobile channels in the point of impact, for example in a store, are really promising.

6.3. Case 3

The third case is described based on the interview of Consumer Analytics Director from a large telecommunications company. The interviewee's department practices consumer analytics based on both to traditional market researches and to consumer data in the company's systems. Additionally, also consumer data based reporting and OLAP are part of the department's tasks. The interviewee itself has been working in the case company for three years now and leads a team which is working especially with advanced analytics. On top of that the interviewee has been working before as a management consultant for over 10 years. Many of the interviewee's previous projects have been related to customers and customer relationship management. Therefore the interviewee's background provides a good view to the business and different analytical methodologies and allows the interviewee to work as intermediate between the business and analysts.

Additionally, also the company has a long background in consumer and marketing research. However, during the last three years there have been needs to understand the consumers even better and therefore analyze the internal consumer related data. Furthermore, also traditional business intelligence such as sales reporting has been practiced for a long time in the company. For example the first data warehousing project has been started over 10 years ago. In addition, also some consumer related reporting has existed for a long time but in the past all this was divided throughout the company. In year 2008 the company started really to centralize its data into one common data warehouse. Before this most of the consumer related data was in operational databases or in departmental data marts.

6.3.1. *Business intelligence architecture*

The company collects detailed consumer data for example from the transactional systems related to its services and their usage. Also company's websites are used as a data source and basically all the hits to the pages are stored into the data warehouse. Because of the current operating model the company is not however at the moment able to collect basic customer data from those consumers who have bought its products. There are however some exceptions as the company is for example able to collect detailed customer data from a small part of product users who have agreed that the company can monitor their actual product and service usage. This monitoring includes basically everything the consumers do with the products. Furthermore, the company collects external data by conducting large amounts of researches or questionnaires which include questions for the consumers regarding the industry. These researches contain also questions regarding the usage of products and services, future trends, and the brand image. Additionally, these researches are targeted so that the company is able to collect also information from those consumers who use their competitor's products and services. On top of this, data is gathered from different call centers and other touchpoints such as the company's web store. All this data is gathered into a centralized data warehouse. On top of this the company uses departmental or functional data marts.

Furthermore, the company is able to create a single view to a specific consumer especially with that consumer data which is related to those consumers who have agreed that the company can monitor their actions. With other consumers it is however harder to create a single view as identifying link between different sources of data might be unavailable. This is caused by the fact that traditionally the consumers do not really have the need to register themselves so that the company could identify them. The only exceptions to this are call center contacts and web store purchases which both require registration. In these cases the company is able to create a better understanding regarding a specific consumer. The company has however also some other means to create a more comprehensive view to the data as the company's products for example have their own identifier. This identifier can be used to create a comprehensive picture from the usage of a specific product. Additionally, the company is at the moment putting effort into acquiring more identified and registered consumers.

Data is integrated into the centralized data warehouse from most of the data sources in a daily basis. Data relevance on the other hand is not considered as in the company's opinion one cannot know what data will be relevant for example after a year and if the data is not collected to the data warehouse then it is in many cases hard or impossible to acquire anymore. The interviewee also stated that the whole idea of data mining is that the analyst does not in the beginning know what data is relevant but instead tries different kind of things. Furthermore, in the interviewee's opinion it is easy to define what the relevant and needed data is for reporting but especially in predictive analytics it is not possible to do such definition beforehand. A lot of effort is however put in to creating common data definitions so that different people throughout the company will understand the data in a similar way.

Data quality is on the other hand seen as an improvement point. There are especially problems in the different touchpoints which collect the data. Sometimes the inputted data is wrong or missing altogether. The interviewee raises this as the main factor which is causing the problems with the quality of data. Furthermore, the company has also multiple different systems which perform similar functions. All these systems have in addition their own data models and therefore it is not a simple task to combine the data into the data warehouse even though the company uses different ETL-tools and data management processes to integrate and clean the data. Some data quality issues can be fixed during these processes but naturally everything cannot be fixed in this phase.

Additionally, the company uses multiple different centralized business intelligence platforms. The reasons for having multiple different business intelligence tools are user preferences and the different capabilities of different tools. All the different platforms use the same centralized data warehouse as a data source among other smaller data marts with specific purposes. The company has also different analytical tools and applications which are mainly used by the interviewee's department. Some of the tools are used to create more traditional analyses and some of the tools are used to create more operational and real-time analyses. Furthermore, the results of different analyses are delivered to the business users in two main ways. High level insight, such as the characteristics of a typical consumer who uses a certain service, is delivered with visual delivery media such as dashboards or Power Points. An operative insight on the other hand is inputted automatically into the business processes. The interviewee defines the latter as operative analytics in which the company is at the moment more and more focusing on. On top of this, some of the company's departments also have analytical applications which mine the data from the web. These applications have not however been in a wide use and therefore the company is at the moment implementing a centralized web mining application which will be used throughout the company.

6.3.2. *Customer insight and sales*

The company is able to support personalization, targeting, predictions, and insight generation with consumer analytics. As an example, the company has created a new segmentation model based on consumer data analysis. The data used in this segmentation is mainly behavioral data as demographical data is not available for all consumers. Additionally, this segmentation model includes among other things product positioning for different segments in different countries. This information can be used by marketing to define what products should be marketed to a specific consumer segment so that the marketing ROI is improved. Furthermore, consumer analytics is also used to support customer relationship management with different targeting campaigns and customer churn prediction based on to the consumer behavior. For example the consumers which are about to churn can be targeted with appropriate actions which keep them from churning and therefore sustain sales. Additionally, the company has also been able to identify different purchase and consumer loyalty drivers which drive the consumer to buy from the company. Both behavioral data and market research has been used to identify these drivers.

The company is also able to predict next best service and product offerings for specific consumers so that the company's sales can be increased. In other words they are able to predict what consumers are likely to buy next. Additionally, store recommendation engine which is able to propose suitable products or services for a specific consumer using the company's web store have been created. This recommendation is based on to the behavioral data gathered from the web store users and increases sales as consumer is recommended with a product that most likely suits for him.

Additionally, the raw click-stream data from the company's websites is also mined with different data mining tools in order to for example identify what kind of user profiles there exist. These analyses can identify for example behavioral patterns such as how often certain type of user comes to the websites. On top of this the company is able to link many of the website users to the products those users own. This is done based on the specific pages the user is viewing or the campaigns the user is joining. The user might also explicitly state the product during the website session. All these analyses enrich the consumer data the company has. The click-stream data is also used in different kinds of dashboards which monitor for example the website usage.

The interviewee states that in the past market research was the main way to gather information about the company’s brand. At the moment the company is however also centrally starting to use social media as a source for consumer insight. This insight can include consumer opinions about the company brand and its products and services. In addition, also some product faults can be noticed through social media before the consumers start to contact the call centers. Additionally, product development uses detailed product usage data which can be collected from those consumers who have agreed to be monitored to further develop products.

Furthermore, the company uses analytics also for different things than customer analytics. These include for example demand-supply and product pricing related analytics. In the future the company will focus especially on conceptual services which can be created with the help of operational and real-time analytics. Location based marketing is an example of such a conceptual service.

6.4. Case 4

The fourth case description is written based on the interview of Information Services Director and eCommerce Development Manager from a large retail company. The Information Services Director is currently working in the case company in a department related to information and analysis services and has also previously worked at a department related to customers and marketing. Altogether the interviewee has worked in the case company for three and half years. The interviewee has however worked with marketing development, performance management, and analytics already for over 10 years in different companies. The eCommerce Development Manager on the other hand has worked in different positions in the company for over 10 years and has therefore good experience from the industry. In the current position related to web store development and web analytics the interviewee has been now for three years. In addition, the interviewee has similar experience also from other companies.

Furthermore, both interviewees have hands-on experience from different analytical applications and tools. They are also both familiar with reporting tools and applications which are used in the company. The company has been practicing reporting and analytics already for multiple years.

6.4.1. *Business intelligence architecture*

There are multiple different data sources in the company. The largest data source is the loyalty program and related systems which contain large quantities of data from individual customer's sales transactions and demographics. Other internal data sources include among other things operational systems, websites and web store. In addition, the company uses many external data sources such as market researches, customer panels and public customer data sources. This allows the company to get important data both from its own customers and from its competitors and their customers. Unstructured data sources are not however currently used in the company.

All the data which is used in reporting and analytics is considered to be quite relevant. The company sees that the data used in decision making needs to be suitable for the decision in hand. Non-relevant data can cause wrong decisions or it can slow down the decision making as it slows the finding of relevant data. Furthermore, the relevant customer data from both internal and external sources is combined into centralized data warehouse so that the company would have a more comprehensive view to the customers. This is done with enterprise-wide scope and the data used in reporting and analytics is integrated into the data warehouse on daily basis with overnight batches. In certain business units this is done even on hourly basis so that performance management and analyses can be used to make fast decisions. Additionally, there is currently a growing trend in the company for fastening the data integrations.

Furthermore, there have been challenges in the company regarding data quality because the data volumes are large and there are multiple data sources. In some cases it is hard to combine the data from different sources while in other cases some data might be missing. Recently the company has however been able to overcome many of these challenges and data quality is currently in quite good shape. For example no wrong decisions have been made because of poor quality data. Instead, the data quality is assured with processes which make sure that data is not used if it cannot be trusted. This has however meant that some of the data cannot be therefore used in reporting or in analytics.

The company uses different analytical technologies for different purposes. There are tools and applications for analyzing market researches, website usage, and marketing results. Also tools and applications which can be used to analyze comprehensive customer data for multiple different purposes such as data mining tools are used. Even though for example social media is seen as interesting source of customer insight, text or web mining tools are not yet in use in the company. Additionally, the different analytical technologies are not really integrated into business processes at the moment. Furthermore, customer analytics is practiced partly with centralized model and partly with decentralized model. Some of the main analyses are done by the centralized information and analysis services team while some other analyses are done by the different departments or business units themselves. This is caused by the fact that some of the analyses are needed in a short notice and their development requires deep business understanding. On top of this, the company sees that in many cases customer analytics and business analytics in general is team work as analyses made purely based on data can be wrong. The analysts need to analyze the data together with the business so that also business understanding is included into the analyses.

In addition, the company also has reporting tools which are used for producing reports and for supporting business performance management. Additionally, functionalities such as dashboards are used. Furthermore, the company uses mainly one centralized reporting tool for general reporting, but also other centralized reporting tools are used. For example reports regarding website and web store usage and marketing effectiveness are produced with specific reporting tools.

6.4.2. *Customer insight and sales*

Overall the company does mainly more descriptive customer analytics which brings insight about the structure and behavior of the customer base instead of practicing predictive analytics. For example segmentation is done based on both demographical and behavioral customer data. This segmentation is used for instance when marketing strategies are defined. Additionally, also marketing can be targeted based on the needs of specific customer segments for those customers who are part of the loyalty program. Furthermore, segmentation is also used when the product offering is planned.

The company is also able to categorize customers anonymously based on their profitability for the company. This information is not however used much as the company has made a strategic alignment that all customers are basically treated in a similar fashion. The only exception to this is the difference between loyalty program customers and other customers. Therefore the profitability information is not used for example to give different level of service or offers for more profitable customers. Furthermore, the company uses also for example market research and website usage data when profiling its customers. This data is additionally combined with the sales transactions so that a more comprehensive customer profiles can be identified. Such combination can be done both with on-line and off-line transactions.

The interviewees state that the customer data, which is collected through customer loyalty program, is the most extensively analyzed data in the company. The company is for example able to analyze the sales transactions, customer segments, and customer behavior based on geographical distribution. As an example, they are able to analyze the impact area of a specific store and use this insight when new store locations are planned. The location with the highest customer potential could be for example selected and therefore the increase of sales would be maximized. Other examples include for instance the identification of unharnessed customer potential in a specific area.

Furthermore, also customer related events are analyzed and identified to some extent in the company. If a customer for example moves to different city or gets a child, he can be approached with targeted marketing which contains offers related to that customer event. Additionally, there have been first demonstrations regarding econometric modeling which identifies those variables which affect to the increase of sales. The idea is to identify the effects of company's marketing related activities versus the effects of for instance weather or season related variables.

On the contrary, changes in customer's behavior are not really followed at the moment. For example customer churn is not currently predicted. One reason for this is the industry characteristics which cause churn prediction to be really complex. In addition, also the loyalty program has been such a success that it prevents automatically customer churn to some degree as the more the customers buy from the company the more benefits they receive. Additionally, also the cross-selling and up-selling possibilities are not currently predicted but the company is currently trying to create analyses which would allow automatic product recommendations based on for example the previous purchases of the customer. Furthermore, web mining is also something what the company will consider in future. Web mining is expected to give a better picture about the customers' attitudes and opinions regarding the company's brand and products. It is also seen as a good source for customer feedback.

All in all, the company sees that there are still many areas where customer insight is needed. Especially marketing processes and customer relationship management should be automated with customer analytics. For example the conversion rates of different marketing channels and campaigns could be analyzed automatically and the best approaches could be then automatically selected. Same applies for targeted marketing. Furthermore, there exist already possibilities to create targeted marketing based on multiple different customer related variables but this is not yet used to its full potential. Also the sales management of the company should be moved from reactive approach to proactive approach. The company should for instance be able to forecast sales more precisely.

6.5. Case 5

The fifth case is described based on the interview of Development Manager - Analytical CRM and Data Mining and Senior Development Manager from a large telecommunications company. The Development Manager - Analytical CRM and Data Mining has worked in the case company for over ten years with business intelligence and customer analytics. Therefore, using for example data mining tools has been part of the interviewee's everyday job. Furthermore, the interviewee has also previous experience related to market research, statistics, and analytics from other companies. The Senior Development Manager on the other hand has worked in various different positions related to business intelligence all the way from year 1995 from which the last ten years in the case company. These positions have included especially tasks related to data warehousing and business intelligence tools.

Both of the interviewees have a good hands-on experience from different tools and applications related to business intelligence or business analytics. In addition, also the case company has a long experience from data warehousing and business intelligence. The first related initiatives have been started around 15 years ago and for instance data mining and customer analytics have been practiced in the company already for over 10 years.

6.5.1. *Business intelligence architecture*

The company's main sources of customer data are the three main internal CRM and billing systems which include the basic customer data, contracts, sold products and services, billing history, and so on. On top of this, also other relevant internal systems are used. These include for example ERP system and company's websites from where the site usage related data is collected. Also some other systems which have data related to customer behavior are used. Additionally, the company also buys data from different external customer data providers. These external data sources include for example different registries. Furthermore, the data from these external sources is connected to the data from internal sources so that more comprehensive customer profiles can be created. This is true for both individual customers and business customers.

In addition to this also market researches are used to gather customer data. The data gathered through market researches is however mainly handled separately as it cannot be connected to individual customers. On top of this, also data related to competitors and their activities is collected in large volumes because there is a tough competition in the telecommunications industry. On the contrary, unstructured data sources such as contracts are not currently brought into any data warehouse. The most important details of those contracts are however also stored in structured format in CRM systems and therefore they can be used in analyses. Additionally, the first initiatives regarding unstructured data in web and social media have been started recently.

Furthermore, the company has currently a few data warehouses. This is mainly due to different mergers and historical reasons which have introduced additional data warehouses. The different data warehouses are however consolidated now and then in order to create more centralized data warehouse architecture. This consolidation process is however slow because in many cases there are no clear business needs to combine especially some of the older data warehouses. In order to be able to create a comprehensive and enterprise-wide view into the customer data the company has however introduced a new data warehouse which includes data from all the main customer data sources.

The data from different source systems is integrated and combined into the data warehouses in batches with an ETL-tool. In most of the cases the data integration is done once a day but in some cases, where there is a specific need for more topical data, the integration is done multiple times in a day. In addition, the ETL-process includes different monitoring mechanisms which identify and report possible data quality issues. These issues are then forwarded to the responsible persons of different sources systems so that the data quality issues could be fixed already in the source systems. Nevertheless, data quality problems are seen as one of the biggest obstacles for business intelligence. In many cases the different data quality problems are noticed for example when reports or analyses are done based on the data. The current assumption is that data is correct in the source systems but it rarely or never actually really is. This is especially true with some of the older source systems. On the contrary, many of the newer systems have exceptionally high data quality.

Furthermore, the company uses one centralized and sophisticated business intelligence tool in two different centralized business intelligence portals. One of the portals is for internal use and the other is used by the company's partners and customers. Basically all the company's reporting is done with these two portals and therefore the business intelligence tool is connected to the different data warehouses. Of course, also operational systems have some reporting or data extract functionalities but they are not considered as business intelligence. Instead this functionality is regarded as a normal functionality of the operational systems. The company has also additionally a centralized and sophisticated analytical tool which is used for multiple different application areas and which is capable for example for data mining and predictive analytics. In addition, the tool can be used to integrate the results of different analyses into the business processes. On top of this, there exist also some other analytical tools, such as web mining tools, but the centralized tool is used for the majority of the analyses.

6.5.2. *Customer insight and sales*

The company practices customer analytics in a wide range of different application areas. Both individual customers and business customers are extensively analyzed. These analyses include among other things churn prediction and management, customer behavior analyses, segmentations or classifications, sales forecasts, and so on.

The company uses multiple customer classifications and segmentations which are created based on a variety of different types of customer data. These customer classifications are done based on different business needs and additionally these classifications are modified often. Especially the industry characteristics have introduced many changes as old technologies and services have been replaced by new ones. Furthermore, these customer classifications are used for example for targeted marketing. In addition to the customer classifications, the company for instance also analyzes the website usage in order to identify different user profiles.

Also customer loyalty has been analyzed and studied so that the company knows who from its customers are loyal and who are not. The company has also identified those factors or elements which affect to the customer loyalty and help in keeping the customers loyal. Furthermore, this information is integrated into the business processes. However, in the interviewees' opinion customer loyalty and different actions related to customer retention are in some cases overrated because in telecommunications industry competitors can so easily react to these different actions.

Furthermore, also customer profitability and customer churn is followed through different analyses. The interviewees see customer profitability as an important part of customer related decision making. They say that it is an essential part of for example churn prediction and management to know if the customer is a profitable or not. The company focuses its efforts especially on those customers who are both profitable and in danger of starting to use competitors' products or services. This information is integrated to the business processes so that correct actions can be performed fast enough.

On top of this, also different customer related events such as change in address or in life situation are monitored, analyzed, and reacted upon. The interviewees see that these events are especially potential moments for example to acquire new customers by marketing appropriate products at the right time. This applies also to those customers who normally do not react to offers. In addition, different cross-sell opportunities are widely identified with multiple different data mining models. These include for example identifying customer needs based on those products that a specific customer is missing compared to other similar customers. Therefore both different product associations and propensity to purchase are analyzed. This information is also integrated as part of the business processes so that these identified opportunities can be utilized. Furthermore, the company is also into some extent using a recommendation engine in their websites which proposes possibly suitable products for the users.

Furthermore, the company also uses analytics to validate the effectiveness of different marketing campaigns. Based on these analyses the effective campaigns are continued or further developed and the non-effective campaigns are closed. Additionally, there have been also experiments to increase sales through social network analysis where the group leaders of social networks are identified. This information cannot however be used to its full potential as this kind of analysis needs to be done unanimously according to the Finnish legislation. The information can however be used to identify those customer segments which most likely contain such group leaders of social networks. These customer segments can be then targeted with specialized marketing. On top of this, the company has started its first small experiments with web mining regarding brand and product management.

All in all, different customer analyses and classifications are usually integrated into the business processes so that the company's employees can react based on to this information whenever they are in contact with customers. This applies to both individual customers and business customers. The interviewees see that such integration is really important part of business intelligence and business analytics. Furthermore they state that the challenge in this is to figure out how the customer insight can be brought to the different customer touchpoints in a systematic and sensible way. In addition, the interviewees see business intelligence and business analytics as a strategic differentiator especially in the telecommunications industry as they allow the company to understand its customers and their behavior.

Furthermore, the company is using analytics and analytical technologies also for many other applications areas besides customer analytics. These areas include among other things product pricing, procurement support, and service improvement. For example the performance and functionality of call centers have been greatly improved with different analyses.

7. RESULTS

The results of this research are introduced in this chapter. Firstly, the business intelligence architecture maturities of the case companies are evaluated based on to the case descriptions in chapter 6 and to the business intelligence architecture maturity model which was composed in chapter 4. Thus, the maturity is evaluated in the following categories: data sourcing, data quality and integration, data warehouse architecture and scope, analytical technologies, and business intelligence technologies. In addition, the maturity is evaluated especially from the perspective of customer analytics.

Furthermore, the role of business intelligence architecture in customer analytics is analyzed in the second section. This is done by comparing the business intelligence architecture maturities of the case companies to the application areas of customer analytics the case companies practice. Second section answers therefore to the following research question:

How does the maturity of different aspects of business intelligence architecture affect the insightfulness of customer analytics?

After this the following main research question is answered by analyzing the results of this study:

Does the use of comprehensive customer information and advanced analytical applications create insight which can be used to further increase sales?

7.1. Business intelligence architecture maturities of case companies

The first case company has a centralized data warehouse which has multiple data sources and is used throughout the enterprise. Therefore the data warehouse architecture and scope category maturity can be seen to be on adult stage. All the maturity levels of the first case company are illustrated in Figure 14.

The used data sources include different internal operational systems, web store and some external partner related data sources. Also market research is performed into some extent but the research data is not connected to the data warehouse. All in all, the company is able to collect a lot of data from business customers and partners but it is not currently able to create a comprehensive view to individual customers. Either the data does not at the moment exist or it is not collected into the data warehouse.

Additionally, unstructured data in the internal documentation or in the internet is not used at the moment at all. Thus, the data sourcing category maturity is on child stage as a lot of possibilities still exist related to the comprehensiveness of customer data. Furthermore, the company mainly integrates data into the data warehouse with overnight batches. In addition, critical data can be integrated during days. Data quality on the other hand is mainly monitored with manual processes even though the company uses data quality checks in the ETL-process and certain master data management tools. As some poor quality data however reaches the end users the data quality and integration category maturity can be defined to be on teenager stage.

In the analytical technologies category maturity the company is still on a quite low stage. The company uses OLAP for all the analyses it does as no other analytical technologies are available. Thus, the maturity can be defined to be on child stage. On the contrary, in the business intelligence technologies category the company is on a higher stage. It has a reporting portal for example and the business intelligence technologies are used widely in the company. The business intelligence tools are not however fully centralized and business intelligence services are not really provided for example for partners. Therefore, the company can is on teenager stage in the business intelligence technologies category maturity.

Stage / Category	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(A) Data sourcing	Not comprehensive		Somewhat comprehensive		Fully comprehensive
(B) Data quality & integration	Non-trustworthy and not timely data		Somewhat trustworthy and timely data		Fully trustworthy and timely data
(C) Data warehouse architecture & scope	Individual spreadmarts	Deparmental non-integrated data marts	Business unit/division level non-integrated data warehouses	Central enterprise data warehouse with or without data marts	Data service via service-oriented architecture with inter-enterprise scope
(D) Analytical technologies	None	Isolated analytic efforts and stand-alone analytical tools	First initiatives to use sophisticated analytics	Sophisticated analytical technologies	Highly sophisticated analytical technologies
(E) Business intelligence technologies	Only operational reports and extracts	Isolated BI efforts and stand-alone reporting tools	Proliferation of BI tools	Centralized BI tools	BI / BA via service-oriented architecture

Figure 14. Business intelligence architecture maturity from customer analytics perspective in case 1.

The second case company on the other hand has a few different data warehouses. These data warehouses are however integrated so that a comprehensive picture to customers throughout the enterprise is available. For this reason the company's maturity stage in the data warehouse architecture and scope category from the perspective of customer analytics is closer to adult than teenager stage. This and all the other maturity stages of the second case company are illustrated in Figure 15.

Furthermore, the company's data warehouses include internal data from a wide range of operational systems and external data from market researches. The company is able to collect a lot of data also from its competitors' customers and therefore create a comprehensive view to both existing and potential customers. However, the customer data could be even more comprehensive as for example structured or unstructured data in the internet is not used. Therefore, the data sourcing category maturity of the company can be seen to be on adult stage. Furthermore, data integration between the data sources and the data warehouse is currently done once a day with batches. Additionally, data quality related processes are currently more and more automated and data quality is considered to be at least sufficient. There are however still many improvement points regarding the data quality and integration. Thus, the maturity of this category is on teenager stage.

The company uses sophisticated analytical technologies which allow the company to perform data mining, predictive analytics, and simulation. These analytical technologies are also integrated into the different business processes. However, the company does not have yet analytical technologies which would allow text or web mining. Therefore the analytical technologies category maturity is on adult stage. Additionally, the company uses a sophisticated centralized business intelligence platform which provides tools for the whole company's reporting. Thus, the maturity of business intelligence technologies category is on adult stage.

Stage / Category	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(A) Data sourcing	Not comprehensive	←→	Somewhat comprehensive	←→	Fully comprehensive
(B) Data quality & integration	Non-trustworthy and not timely data	←→	Somewhat trustworthy and timely data	←→	Fully trustworthy and timely data
(C) Data warehouse architecture & scope	Individual spreadmarts	Deparmental non-integrated data marts	Business unit/division level non-integrated data warehouses	Central enterprise data warehouse with or without data marts	Data service via service-oriented architecture with inter-enterprise scope
(D) Analytical technologies	None	Isolated analytic efforts and stand-alone analytical tools	First initiatives to use sophisticated analytics	Sophisticated analytical technologies	Highly sophisticated analytical technologies
(E) Business intelligence technologies	Only operational reports and extracts	Isolated BI efforts and stand-alone reporting tools	Proliferation of BI tools	Centralized BI tools	BI / BA via service-oriented architecture

Figure 15. Business intelligence architecture maturity from customer analytics perspective in case 2.

Furthermore, the third case company collects customer data from multiple internal and external data sources. These include a variety of internal operational systems, company's websites, consumers itself, market researches, and social media. Even though the customer data is quite comprehensive there are however things to improve especially regarding the demographical data of those consumers who have bought the company's products. Therefore the data sourcing category maturity can be defined to be on adult stage. This maturity stage and all the other maturity stages of the third case company are illustrated in Figure 16.

The company has an enterprise-wide centralized data warehouse with data marts. Therefore the company's maturity stage in the data warehouse architecture and scope category can be seen to be on adult stage. Furthermore, the data from data sources is integrated into the data warehouse usually once per day. Data quality of this data is on the other hand not as good as the company would like. Even though the company uses ETL-tools and data management processes to improve data quality all the quality problems, which originate from the data sources, cannot be fixed. Thus, the data quality and integration category maturity is on teenager stage.

The company's business intelligence technologies category maturity is on adult stage as the company has multiple centralized business intelligence platforms which provide a wide variety of reporting functionality. Additionally, the company uses different highly sophisticated analytical technologies which provide functionality among other things for data mining, predictive analytics, web and text mining, and operative analytics. Therefore the analytical technologies category maturity is on sage stage.

Stage / Category	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(A) Data sourcing	Not comprehensive		Somewhat comprehensive		Fully comprehensive
(B) Data quality & integration	Non-trustworthy and not timely data		Somewhat trustworthy and timely data		Fully trustworthy and timely data
(C) Data warehouse architecture & scope	Individual spreadmarts	Deparmental non-integrated data marts	Business unit/division level non-integrated data warehouses	Central enterprise data warehouse with or without data marts	Data service via service-oriented architecture with inter-enterprise scope
(D) Analytical technologies	None	Isolated analytic efforts and stand-alone analytical tools	First initiatives to use sophisticated analytics	Sophisticated analytical technologies	Highly sophisticated analytical technologies
(E) Business intelligence technologies	Only operational reports and extracts	Isolated BI efforts and stand-alone reporting tools	Proliferation of BI tools	Centralized BI tools	BI / BA via service-oriented architecture

Figure 16. Business intelligence architecture maturity from customer analytics perspective in case 3.

The fourth case company on the other hand has a centralized data warehouse where customer data from different data sources is combined so that a comprehensive view to customer data throughout the company is available. Therefore the data warehouse architecture and scope category maturity of the company is on adult stage as illustrated in Figure 17. The figure includes also all the other maturity stages of the fourth case company.

The data sources include multiple different external and internal data sources. Internal data sources include for example a variety of operational systems and websites. External data sources on the other hand include different market researches and public customer data sources excluding unstructured data in social media and other such places. Thus, the data sourcing category maturity can be seen to be on adult stage. Same applies for the data quality and integration category maturity as most of the data quality related issues have been fixed with data quality related processes. In addition, the data integration is done even on hourly basis if needed and therefore the data is quite timely.

The company uses additionally many different analytical technologies which allow the company to use different data mining techniques. However, these tools are mainly not integrated to business processes and they are mainly only used to identify patterns from the analyzed data. Therefore, the analytical technologies category maturity can be seen to be on the far end of teenager stage as some improvements are still needed in order to reach adult stage. The company also uses business intelligence tools for performance management and reporting. There exists multiple different tools but all the tools are centralized. Thus, the maturity of business intelligence technologies category is on adult stage.

Stage / Category	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(A) Data sourcing	Not comprehensive	↔	Somewhat comprehensive	↔	Fully comprehensive
(B) Data quality & integration	Non-trustworthy and not timely data	↔	Somewhat trustworthy and timely data	↔	Fully trustworthy and timely data
(C) Data warehouse architecture & scope	Individual spreadmarts	Deparmental non-integrated data marts	Business unit/division level non-integrated data warehouses	Central enterprise data warehouse with or without data marts	Data service via service-oriented architecture with inter-enterprise scope
(D) Analytical technologies	None	Isolated analytic efforts and stand-alone analytical tools	First initiatives to use sophisticated analytics	Sophisticated analytical technologies	Highly sophisticated analytical technologies
(E) Business intelligence technologies	Only operational reports and extracts	Isolated BI efforts and stand-alone reporting tools	Proliferation of BI tools	Centralized BI tools	BI / BA via service-oriented architecture

Figure 17. Business intelligence architecture maturity from customer analytics perspective in case 4.

Furthermore, the fifth case company has currently a few different data warehouses but one of the data warehouses however contains more comprehensive data from multiple smaller data warehouses. Also customer data is combined from different smaller data warehouses so that a comprehensive and enterprise-wide view to both individual and business customer data is acquired. For this reason the data warehouse architecture and scope category maturity can be seen to be on adult stage especially from the perspective of customer analytics. This and all the other maturity stages of the fifth case company are illustrated in Figure 18.

The company collects data from multiple different internal and external data sources. The internal data sources include for example different CRM and billing systems, ERP systems, websites and systems which include behavioral customer data. External data sources on the other hand include market researches, external data providers, and so on. Also first initiatives to collect customer data from web and social media have been started. All in all, data sourcing is really comprehensive and basically all the different possible customer data sources are used. Therefore the data sourcing category maturity for the company is on sage stage. Data integration is done considering the business needs. Most of the data is integrated once a day and the rest of the data multiple times in a day. Data quality on the other hand is seen as an obstacle as part of the data has quality problems. Data quality is however monitored during the ETL-processes but there are not sophisticated data quality related tools in place. Because data quality issues are affecting business intelligence and business analytics to some degree the data quality and integration category maturity can be seen to be on teenager stage.

On the contrary, the company has a sophisticated centralized business intelligence tool which is used for the whole company's reporting. Thus, the maturity of business intelligence technologies category is on adult stage. On top of this, the company also uses a highly sophisticated analytical tool which allows the company to perform data mining, predictive analytics, and so on. This analytical tool is also integrated into the different business processes. Additionally, the company uses also analytical technologies which allow web mining. Therefore the analytical technologies category maturity can be seen to be on sage stage.

Stage / Category	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(A) Data sourcing	Not comprehensive	←→	Somewhat comprehensive	←→	Fully comprehensive
(B) Data quality & integration	Non-trustworthy and not timely data	←→	Somewhat trustworthy and timely data	←→	Fully trustworthy and timely data
(C) Data warehouse architecture & scope	Individual spreadmarts	Deparmental non-integrated data marts	Business unit/division level non-integrated data warehouses	Central enterprise data warehouse with or without data marts	Data service via service-oriented architecture with inter-enterprise scope
(D) Analytical technologies	None	Isolated analytic efforts and stand-alone analytical tools	First initiatives to use sophisticated analytics	Sophisticated analytical technologies	Highly sophisticated analytical technologies
(E) Business intelligence technologies	Only operational reports and extracts	Isolated BI efforts and stand-alone reporting tools	Proliferation of BI tools	Centralized BI tools	BI / BA via service-oriented architecture

Figure 18. Business intelligence architecture maturity from customer analytics perspective in case 5.

When the business intelligence architecture maturities of the case companies are compared between each other, the differences in the maturities are easily found. This comparison is described in Table 3. The fifth case company has the most comprehensive (5 - sage stage) and the second case company the least comprehensive (2 - child stage) customer data sourcing. The fourth case company is on the highest stage (4 - adult stage) on the data quality and integration category while all the other case companies are in same lower stage (3 - teenager stage). The only category in which all the case companies are in the same stage (4 - adult stage) is the data warehouse architecture and scope category.

Table 3. The business intelligence architecture maturities of all the case companies.

Category	Case 1	Case 2	Case 3	Case 4	Case 5
Data sourcing	2	4	4	4	5
Data quality & integration	3	3	3	4	3
Data warehouse architecture & scope	4	4	4	4	4
Analytical technologies	2	4	5	3	5
Business intelligence technologies	3	4	4	4	4
Business intelligence architecture	**2,8**	**3,8**	**4,0**	**3,8**	**4,2**

Additionally, the third and fifth case companies have the most sophisticated (5 - sage stage) analytical technologies while the first case company has the least sophisticated (2 - child stage) analytical technologies. And lastly, the first case company is in a lower stage (3 - teenager stage) in business intelligence technologies category than the other case companies which are all in the same stage (4 - adult stage). All in all, when the average business intelligence architecture maturity is calculated, the first case company has the lowest maturity stage (2,8). Then comes the second and fourth case companies which have the same average business intelligence architecture maturity (3,8). The fifth case company has the highest (4,2) and the third case company has the second highest (4,0) average business intelligence architecture maturity. The differences in the maturity stages of the different companies are also illustrated in Figure 19. Additionally, a different comparison can be found from appendix 3.

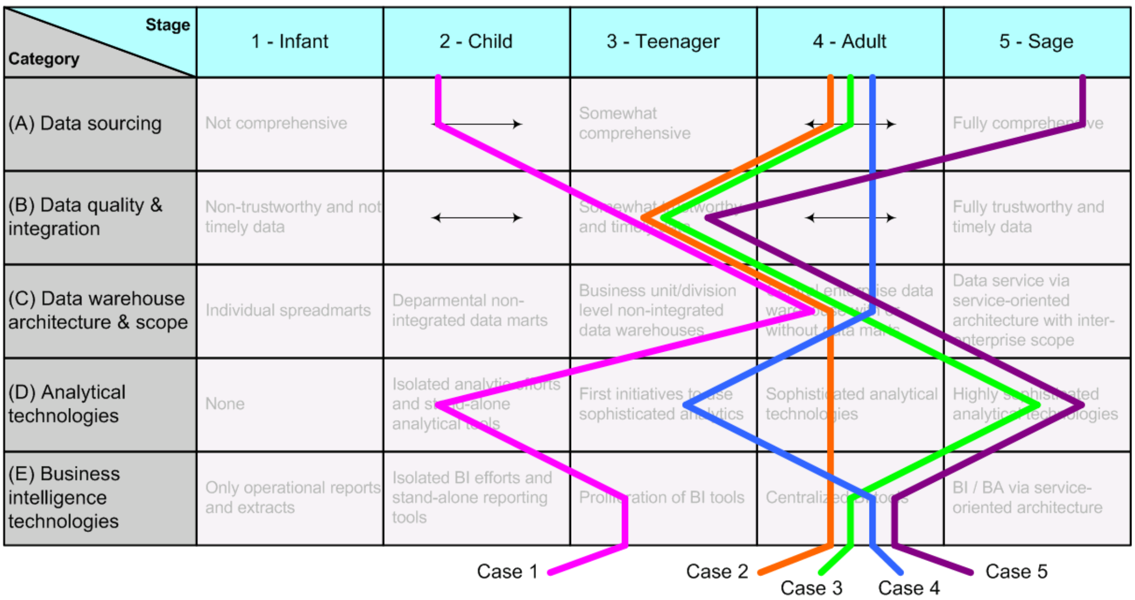

Figure 19. A comparison between the business intelligence architecture maturities of the case companies.

7.2. The role of business intelligence architecture in customer analytics

Each of the case companies practiced customer analytics at least on some level. However, both the amount of different application areas of customer analytics and the level of sophistication vary between the case companies. All the different application areas of customer analytics for all of the case companies are described in Table 4. This listing includes all the application areas which the case companies practiced at least on some level. Many of the same application areas of customer analytics, which were described based on the literature review in chapter 3, are practiced in the case companies. Therefore the application areas are divided into the same introduced categories: customer identification and attraction, customer retention, customer development, and other application areas. Also some additional application areas of business analytics which the interviewees mentioned are included into the table. These however will not be analyzed as some of the interviewees just wanted to mention some, not all, additional areas where business analytics is practiced.

Table 4. The different application areas of customer analytics in the case companies.

Application area of customer analytics	Case 1	Case 2	Case 3	Case 4	Case 5
Customer identification & attraction					
Sophisticated segmentation	Yes	Yes	Yes	Yes	Yes
Customer profitability	Yes			Yes	Yes
Customer profiles		Yes	Yes	Yes	Yes
Targeting	Yes	Yes	Yes	Yes	Yes
Customer retention					
Customer loyalty	Yes	Yes	Yes		Yes
Churn prediction		Yes	Yes		Yes
Customer events		Yes		Yes	Yes
Customer development					
Customer lifetime value		Yes			
Product associations		Yes	Yes		Yes
Customer needs		Yes	Yes	Yes	Yes
Propensity to purchase		Yes	Yes		Yes
Customer recommendations			Yes		Yes
Social network analysis					Yes
Other application areas of customer analytics					
Brand & product management			Yes		Yes
Product development based on usage			Yes		
Sales forecasting	Yes	Yes	Yes	Yes	Yes
Marketing campaign analytics				Yes	Yes
Count:	**5**	**11**	**12**	**8**	**15**
Additional mentioned application areas of business analytics					
Fraud detection		Yes			
Supply chain optimization		Yes			
Product prising			Yes		Yes
Service improvement					Yes
Procurement support					Yes

When only the amount of different application areas of customer analytics is analyzed it can be stated that the first case company practices the smallest amount (5) of different application areas. These application areas include segmentation of customers and the identification of customer probability and loyalty. The fifth case company on the other hand practices the largest amount (15) of different application areas of customer analytics. These application areas include among other things a wide variety of descriptive and predictive analyses such as sophisticated segmentation, targeted marketing, prediction of customer related events, needs, and propensity to purchase, identification of product associations, and so on. Furthermore, second, third and fourth case companies are situated between the two extremes. The fourth case company is practicing mainly descriptive analytics which is used to understand the customer base better. Both the second and the third case companies are practicing a variety of descriptive and predictive analytics.

More interesting findings can be found when the business intelligence architecture maturities of the case companies are compared with the amount of different application areas each case company uses. This comparison is done in Figure 20. It seems that there is a link at least at some level between the business intelligence architecture maturity and the amount of application areas of customer analytics as the case companies with a lower business intelligence architecture maturity are able to practice a smaller amount of application areas of customer analytics and vice versa. The only exception is the difference between second and fourth case companies which have the same business intelligence architecture maturity but the amount of application areas of customer analytics varies somewhat. Second case company practices 11 different application areas of customer analytics while the fourth practices eight different application areas of customer analytics.

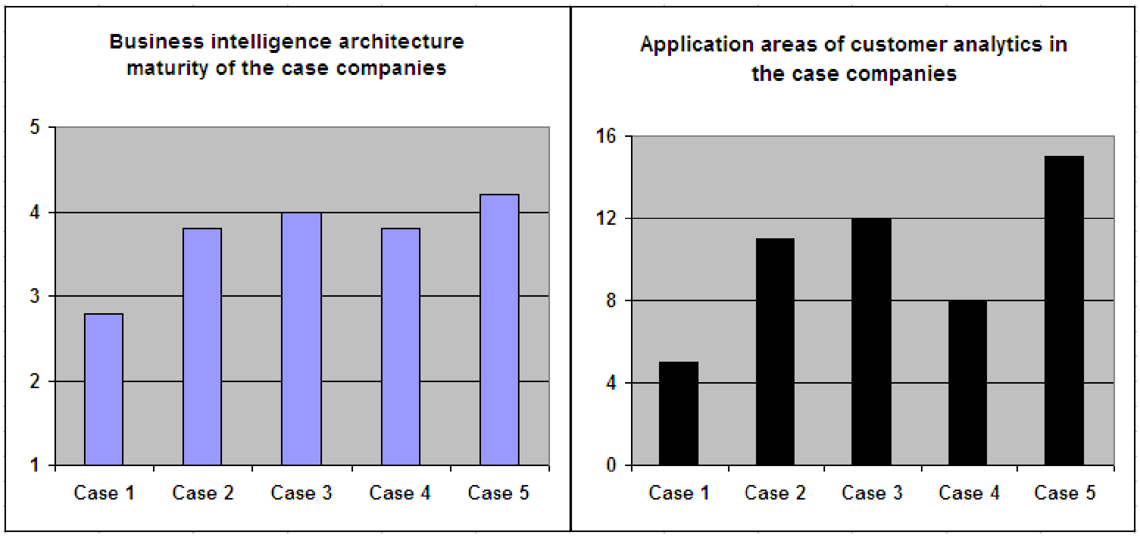

Figure 20. A comparison between business intelligence architecture maturity and the amount of application areas of customer analytics.

In order to understand the possible link between business intelligence architecture maturity and the amount of application areas of customer analytics better, it is beneficial to break the business intelligence architecture maturity back into category maturities. In Figure 21 all the different categories of business intelligence architecture maturity are shown for all the case companies. To ease up the comparison also the amount of application areas of customer analytics for all of the case companies are included.

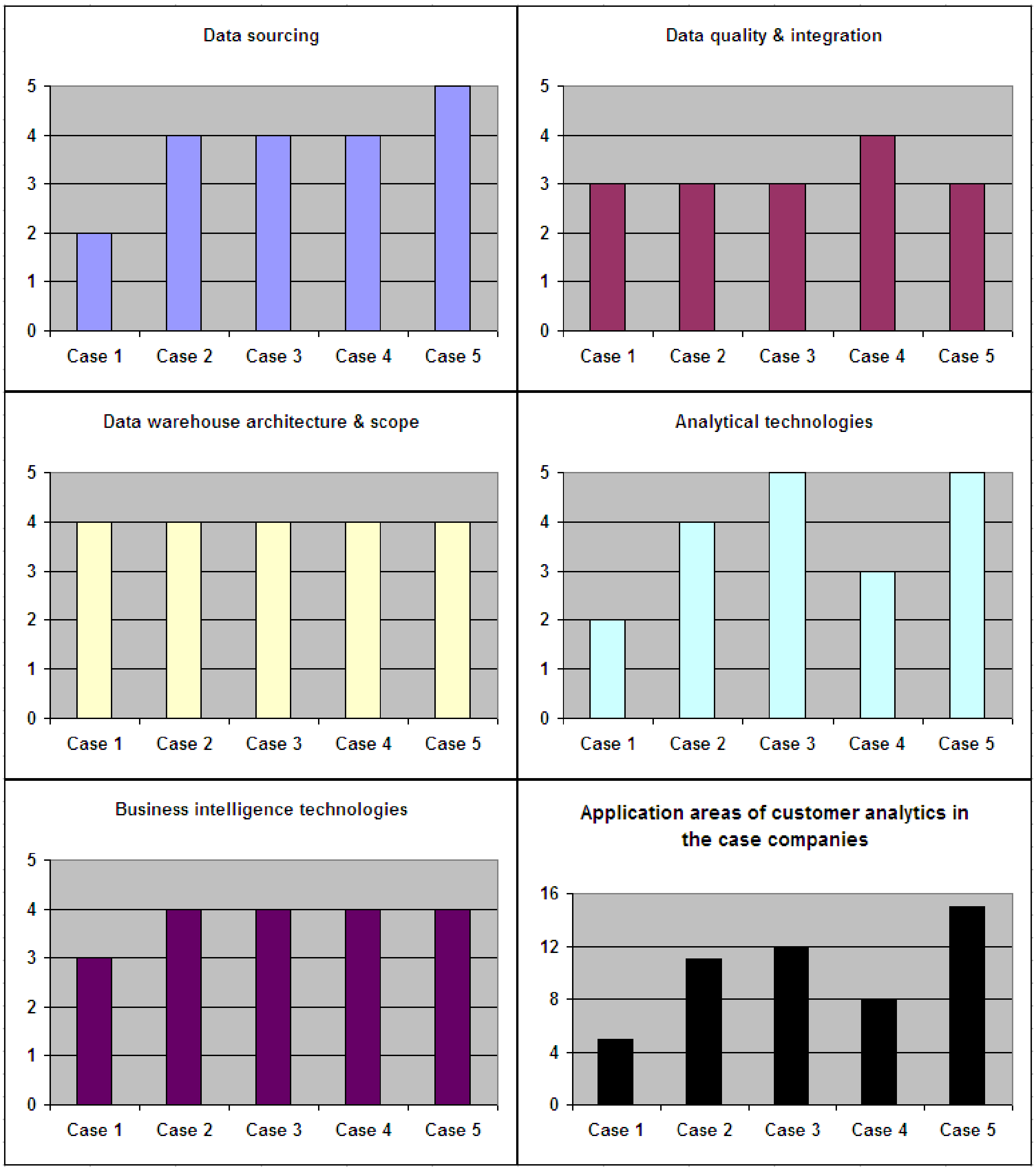

Figure 21. A comparison between the different categories of business intelligence architecture maturity and the amount of application areas of customer analytics.

According to the illustration, the data sourcing maturity seems to follow a similar pattern as the amount of application areas of customer analytics. The first case company has both the lowest data sourcing maturity (2 - child) and the lowest amount of application areas of customer analytics (5) while the fifth case company has the highest values in both. However, the second, third, and fourth case companies have the same category maturities but different amount of application areas of customer analytics. The data quality and integration maturity on the other hand do not seem to be largely linked to the amount of application areas of customer analytics as for example the fourth case company is on the highest maturity stage even though it has the second least application areas of customer analytics. Same seems to apply to the data warehouse architecture and scope maturity as all the case companies are on the same maturity stage even though there are large differences in the amount of application areas of customer analytics.

The analytical technologies maturity however seems to also follow similar pattern as the amount of application areas of customer analytics. The case companies can be ranked based on to the analytical technologies maturity exactly in the same order as based on to the application areas of customer analytics with the exception of third and fifth case companies. They both are on sage maturity stage (5) even though the fifth case company practices more application areas of customer analytics than third case company. On the other hand, the business intelligence technologies maturity does not follow the same pattern with the same accuracy. The first case company is both on the lowest stage of maturity and has the smallest amount of application areas of customer analytics. However, the other case companies are on the same maturity stage even though, as stated already before, there are differences in the amounts of application areas of customer analytics.

The data sourcing and the analytical technologies category maturities seem to have the biggest affect on to the amount of application areas of customer analytics. For this reason the average maturity of these two categories is compared to the amount of application areas of customer analytics in Figure 22.

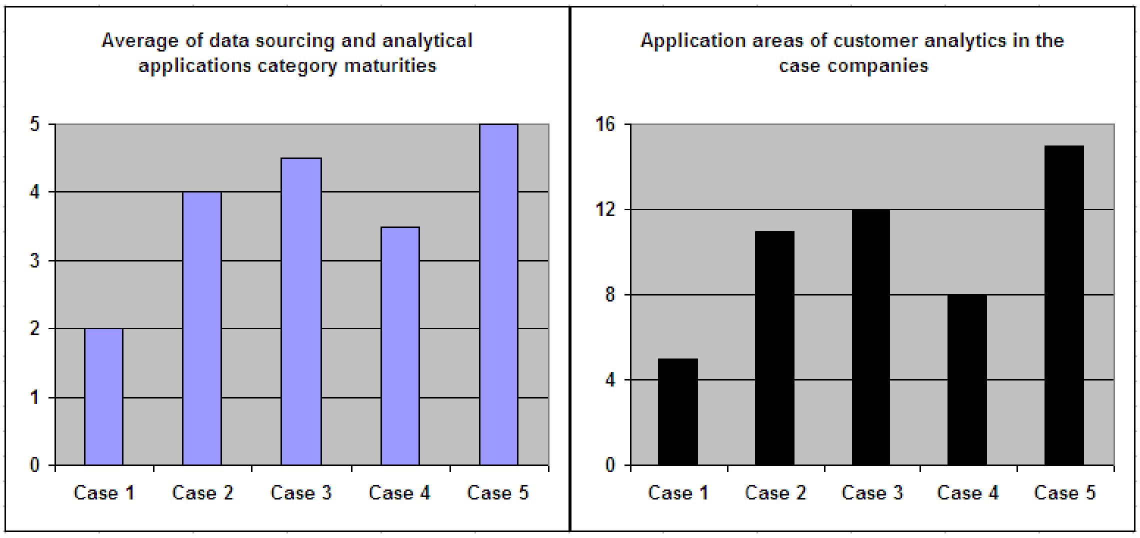

Figure 22. A comparison between the average maturity of data sourcing and analytical technologies categories and the amount of application areas of customer analytics.

Based on this average maturity the case companies can be ranked to the exactly same order as with amount of application areas. It seems quite natural that for example the third case company is able to practice smaller amount of application areas of customer analytics than the fifth case company because it is on a lower stage on data sourcing category even though they are both on the sage stage (5) in analytical technologies category. This is understandable because the third case company possesses less comprehensive customer data which can be used in customer analytics. Same seems to apply also the other way around. The second, third, and fourth case companies all are on the same stage on data sourcing category but there are large differences in the analytical technologies category. In other words the case companies have similar customer data available but the sophistication of their business intelligence architecture regarding analytical technologies is not on same level. Case companies which have more sophisticated business intelligence architecture regarding analytical technologies are able to create different analyses and create more insight from the similar customer data.

As already shortly stated also the level of sophistication varies between the different case companies. The level of sophistication is analyzed for each application area of customer analytics in each of the case companies in Table 5. This analyze values the level of sophistication based on the comprehensiveness of the different application areas and if the analyses are integrated as part of business processes. Naturally if the case company does not practice a specific application area of customer analytics at all, also the level of sophistication is considered to be zero (0). Such a company does gut-based decisions regarding the application area or does not consider the subject at all. One example of such case is the first case company which does not predict when and what products a certain business customer needs but instead sees such prediction as one of the core competencies of the business. On the other hand, if the case company has started to practice a specific application area of customer analytics in part of its business units for instance and uses the obtained customer insight at least to some degree in its decision making, the level of sophistication is considered to be one (1). A good example of such case is the brand and product management of third and fifth case companies. The both companies have just started their first initiatives to use web mining in order to use social media as a source of brand and product related customer insight.

Furthermore, if a specific application area of customer analytics is practiced comprehensively throughout the case company, the level of sophistication is considered to be two (2). The same applies also if a specific application area of customer analytics is practiced at least to some degree and the customer insight is integrated as part of the company’s business processes. For example the fifth case company is at least to some extent able to identify what product or service a specific customer most likely would like to have based on his behavior in the company websites and thus recommend that product or service automatically to the customer. Lastly, the level of sophistication is considered to be three (3) if the case company practices a specific application area of customer analytics comprehensively and has integrated the obtained customer insight as part of its business processes. Great examples from a level three sophistication are the churn prediction and customer profitability related analytics of the fifth case company. The case company is able to identify comprehensively both the profitability of its customers and those customers who are in danger of going to the competitors. On top of this it is able to combine this insight together and integrate it to the business processes so that appropriate actions can be done fast enough. Therefore it can be stated that the higher the level of customer analytics' sophistication is the more customer insight is created and acted upon. In addition, it can be stated that this analysis follows the same principles as the different information management related processes introduces in chapter 2. Information needs to be further processed, distributed to where it matters, and acted upon.

Table 5. The level of sophistication in the different application areas of customer analytics in the case companies.

0 = None or gut-based 1 = Information-based	2 = Comprehensive or integrated into business processes 3 = Comprehensive and integrated into business processes				
Application area of customer analytics	**Case 1**	**Case 2**	**Case 3**	**Case 4**	**Case 5**
Customer identification & attraction					
Sophisticated segmentation	1	3	3	3	3
Customer profitability	1	0	0	2	3
Customer profiles	0	3	3	2	2
Targeting	1	2	2	2	2
Customer retention					
Customer loyalty	1	2	2	0	3
Churn prediction	0	2	3	0	3
Customer events	0	3	0	2	3
Customer development					
Customer lifetime value	0	1	0	0	0
Product associations	0	2	3	0	3
Customer needs	0	2	3	2	3
Propensity to purchase	0	2	3	0	3
Customer recommendations	0	0	3	0	2
Social network analysis	0	0	0	0	2
Other application areas of customer analytics					
Brand & product management	0	0	1	0	1
Product development based on usage	0	0	2	0	0
Sales forecasting	1	2	1	1	1
Marketing campaign analytics	0	0	0	1	3
Total:	**5**	**24**	**29**	**15**	**37**

When the results of the analysis are calculated together it is obvious that there are large differences in the level of customer analytics' sophistication between the case companies. The first case company has the lowest (5) level of customer analytics' sophistication while the fifth company has the highest (37) level of customer analytics' sophistication. All in all, the case companies can be ranked to the same order as with the amount of application areas of customer analytics. These two are compared to each other in Figure 23. Even though the amount of application areas of customer analytics and the level of customer analytics' sophistication seem to follow the same pattern it is important to notice that the differences between the case companies increase when the level of customer analytics' sophistication is considered. This seems to be especially true for the case companies with lower level of customer analytics' sophistication. For example the difference between first case company and other case companies has increased significantly. Also the smaller differences between the second, third, and fourth case companies have increased.

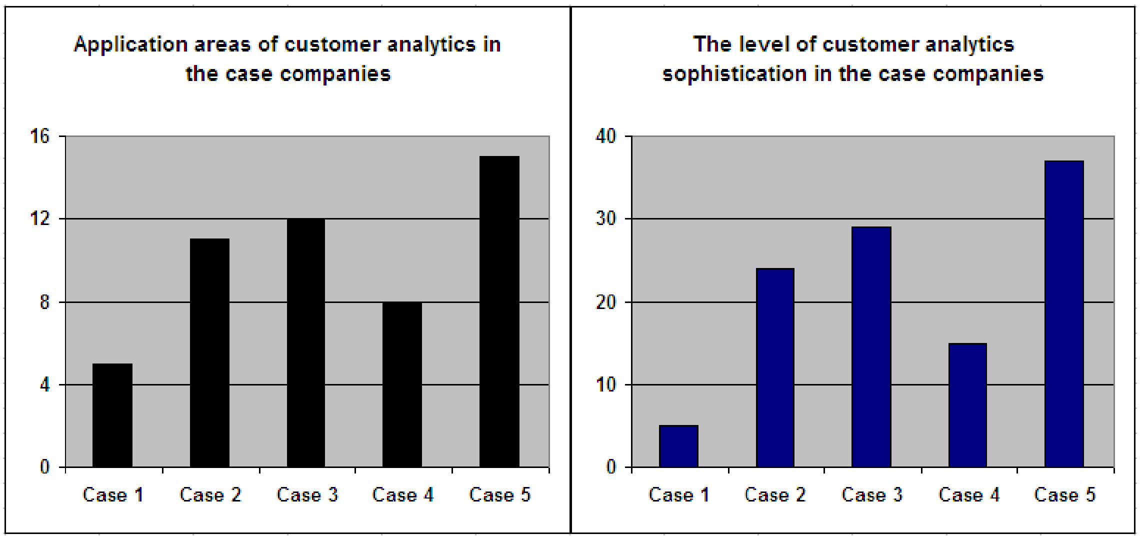

Figure 23. A comparison between the amount of application areas of customer analytics and the level of customer analytics' sophistication.

Furthermore, when the level of customer analytics' sophistication is compared to the business intelligence architecture maturities of the case companies it can be noticed that they both follow the same pattern but there are also large differences. This comparison is illustrated in Figure 24.

The differences between the business intelligence architecture maturities of the case companies are relevantly smaller compared to the differences in the levels of customer analytics' sophistication. This can be easily noticed for example by comparing the differences between first and fifth case companies both in business intelligence architecture maturity and in the level of customer analytics' sophistication. In the latter case the difference is considerably larger. Additionally, there seems to be a large difference between the second and fourth case companies in the level of customer analytics' sophistication even though they have exactly same business intelligence architecture maturity. Especially for this reason it is interesting to compare some of the different categories of the business intelligence architecture maturity into the level of customer analytics' sophistication.

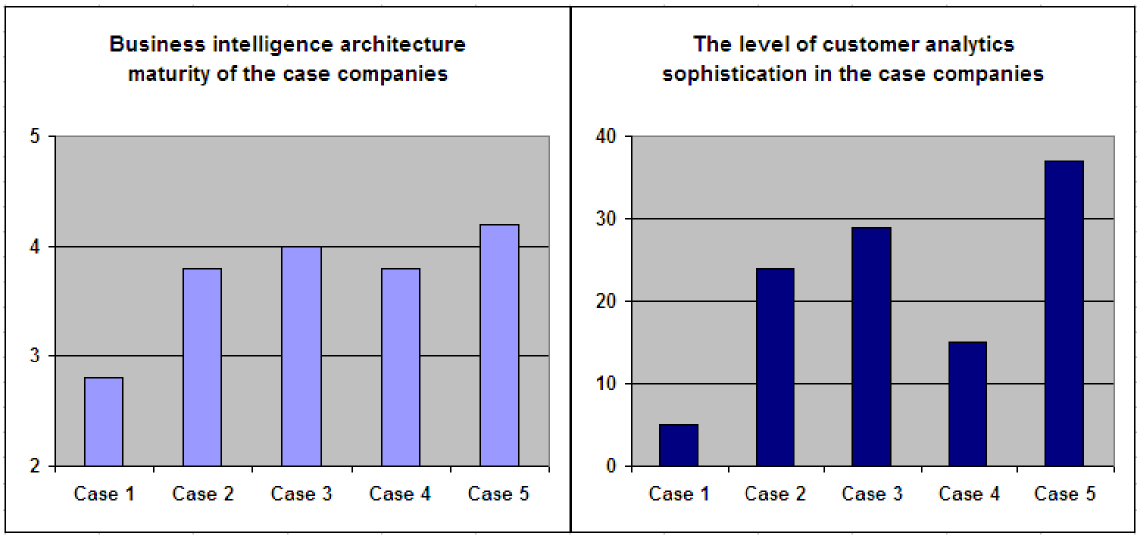

Figure 24. A comparison between business intelligence architecture maturity and the level of customer analytics' sophistication.

However, previously when the amount of application areas of customer analytics was discussed, it was noticed that the data quality and integration, data warehouse architecture and scope, and business intelligence technologies maturities do not seem to follow the same pattern as the amount of application areas of customer analytics. The exactly same situation for the same reasons seems to apply when those categories of business intelligence architecture maturity are compared to the level of customer analytics' sophistication. Therefore the data sourcing and analytical technologies maturities are focused on. These two category maturities and their average maturity are compared to the level of customer analytics' sophistication in Figure 25.

The data sourcing maturity follows to some extent a similar pattern as the level of customer analytics' sophistication. This is noticeable especially from the first and fifth case companies. The first case company has the lowest data sourcing maturity (2 - child) and also the lowest level of customer analytics' sophistication (5). At the same time the fifth case company has both the highest data sourcing maturity (5 - sage) and the highest level of customer analytics' sophistication (37). On the contrary, the second, third, and fourth case companies have exactly same category maturities but the level of customer analytics' sophistication varies a lot between the case companies. Therefore the data sourcing maturity does not seem to fully explain at least alone the differences in the level of customer analytics' sophistication between the case companies.

Additionally also the analytical technologies maturity seems to follow a similar pattern as the level of customer analytics' sophistication. And it actually does it more precisely than the data sourcing maturity. In fact, this is quite natural as the analytical technologies maturity also considers if the analytical applications and tools of the company are integrated as a part of the business processes and therefore deliver the created customer insight to where it can be acted upon. Basically all the case companies can be ranked to the same order based on the analytical technologies maturity as based on to the level of customer analytics' sophistication. The only exceptions are the third and fifth case companies which are on the same analytical technologies maturity stage (5 - sage) even though the fifth case company practices quite a lot more sophisticated customer analytics (37) than the third case company (29). The third and fifth case companies however have different data sourcing category maturities which could explain at least partly the difference in the level of customer analytics' sophistication.

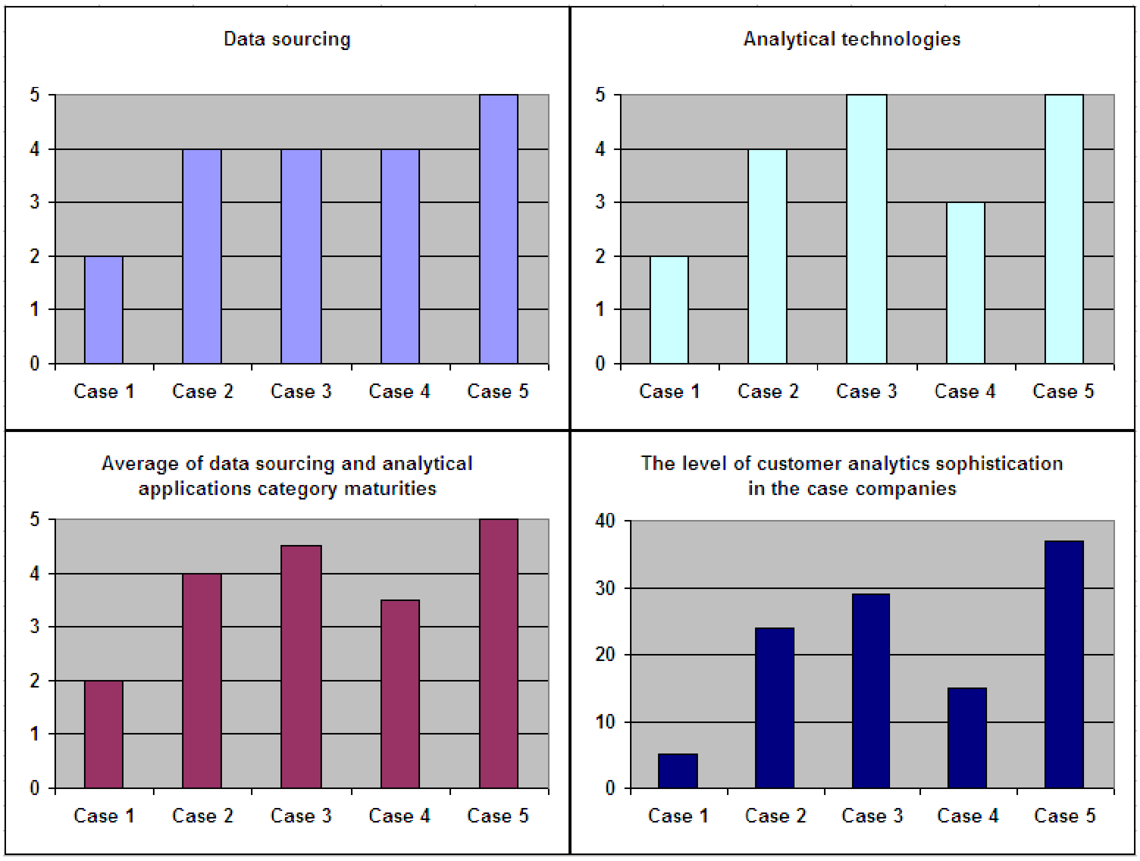

Figure 25. A comparison between data sourcing and analytical technologies maturities, their average maturity and the level of customer analytics' sophistication.

When the average maturity of data sourcing and analytical applications category maturities is compared to the level of customer analytics' sophistication, it can be seen that there really seems to be some level of correlation between the average maturity and the level of customer analytics' sophistication. The higher the average maturity of data sourcing and analytical technologies categories is the higher the level of customer analytics' sophistication is. Therefore both the data sourcing and the analytical technologies category maturities seem to affect to the amount and the usability of created customer insight. In addition, it can be stated that it seems that the infant stage seems to really be a prerequisite which alone is not sufficient to support customer analytics that much. For instance the first case company is on a child stage in both data sourcing and analytical technologies category maturities and is still not able to practice very insightful customer analytics.

All in all, especially the data sourcing and the analytical technologies maturities seem to be linked to customer analytics. Comprehensive customer data, which is collected from a variety of internal and external data sources, makes it possible to do larger amount of different analyses and therefore also affects to the level of company's overall customer analytics' sophistication. Additionally, highly sophisticated analytical technologies, which are integrated into the business processes, make it possible to execute a larger variety of different analyses and especially allow the companies to fully utilize the created customer insight by making it available where and when it matters. However, one has to remember that there exist also other factors, such as the skills of the company's analysts, sponsorship, culture, funding, development processes, organizational objectives, and so on, affecting the insightfulness of customer analytics. All these other factors can also be on different maturity stages in the different case companies and therefore affect the insightfulness of customer analytics in the case companies.

7.3. Business intelligence architecture maturity and the ability to increase sales

In the previous section it was illustrated that there is a correlation between the business intelligence architecture maturity – especially the data sourcing and analytical technologies categories – and the insightfulness of customer analytics. Next the correlation between business intelligence architecture and the ability to increase sales is discussed. As stated in chapter 3 many of the different application areas of customer analytics are related to acquiring new customers, retaining current customers and developing current customers. In other words, many of these application areas of customer analytics are related to increasing sales and sustaining current sales.

Those application areas of customer analytics which are related to customer identification are mainly used to understand the customer base better. Sophisticated segmentation, identification of customer profiles and calculation of customer profitability help companies for example to understand what kind of customers there exist and what kind of customers are the company should try keep or acquire more. For instance first and fifth case companies use customer profitability to target their sales activities. Customer attraction related application areas of customer analytics on the other hand use the understanding which was acquired with customer identification related application areas of customer analytics to attract new and existing customer and therefore increase sales. For example all the case companies except the first case company targets their marketing at least on some level based on their customer segmentation or classification in order to reach the individual customers with more suitable marketing.

Furthermore, the application areas of customer analytics which are related to customer retention are used for example for understanding those factors which keep customers loyal, for analyzing customer events, and for predicting those customers which are in danger of leaving the company. This information can be then used to target company's efforts in order to increase customer retention. Thus, also the current sales can be sustained. For instance the second, fourth, and fifth case companies analyze customer related events for identifying those customers who should be approached with marketing activities in order to either increase customer retention or to harness the possible new sales opportunities.

Moreover, customer development related application areas of customer analytics are used for understanding the different existing possibilities to increase sales. As said, this includes such elements as understanding different product associations, customer needs, social networks, and customer lifetime value. For example the third and fifth case companies are to some extent analyzing customer behavior in their websites or web stores in order to understand customer needs and product associations. This insight is used to give the customers personalized product or service recommendations and therefore increase sales. Additionally different application areas of customer analytics can be for instance used for understanding and developing company's brand, products, and services which in the long run should increase the popularity of the company and its products and services among customers.

For the above mentioned reasons there seems to be a correlation between the insightfulness of company's customer analytics and the company's ability to increase and sustain sales. On top of this there is a correlation between the business intelligence architecture maturity and the insightfulness of customer analytics as discussed in the previous section. This is especially true for the data sourcing and analytical technologies categories of business intelligence architecture. Therefore it can be stated that it seems that there really is a correlation between the comprehensiveness of customer information, the sophistication of analytical applications and the company's ability to increase and sustain sales. Thus, as an answer to the main research question, the use of comprehensive customer information and advanced analytical applications seems to create insight which can be used to further increase sales.

7.4. Assessment of empirical research

Yin (2003, p. 34) states that empirical research can be assessed with four tests. These tests include construct validity, internal validity, external validity, and reliability. Construct validity test defines if the studied concepts are measured with correct measures. Internal validity on the other hand tests the evaluations done in the research regarding the causal relationships between certain conditions. (Yin 2003, p. 34.) Yin (2003, p. 36) argues that the researcher needs to evaluate all the factors affecting certain conditions.

Furthermore, external validity test examines if the research's findings can be generalized (Yin 2003, p. 34). Yin (2003, p. 37) states that usually in a case study the researcher tries to generalize the results of the research into some broader theory. On the contrary, reliability test considers the repeatability of the research. In order to make it possible that another researcher could repeat the exactly same study the researcher needs to document the procedures followed while making the case study carefully. (Yin 2003, p. 37-38.)

When discussing the construct validity, it can be stated that this research used the frameworks related to the business intelligence architecture maturity and to the application areas of customer analytics to measure the studied concepts in the case companies. The same frameworks were also used when the data was collected with semi-structured interviews. This assured the fact that the correct concepts were discussed in the interviews and that the case descriptions, which were composed based on the interviews, also use the same concepts. Additionally, it can be stated that the main concepts of this research such as business intelligence architecture, customer analytics, customer insight, and so on are used in a consistent way throughout the research.

It is hard to evaluate the internal validity of this research as the causal relationships, related to the insightfulness of customer analytics and to the increase of sales, are highly complex. In this research only the causal relationship between business intelligence architecture and these two subjects was studied. Also other factors, affecting especially to the insightfulness of customer analytics, were discussed when the possible causal relationships between business intelligence architecture maturity and the insightfulness of customer analytics were introduced. Additionally, the research does not include such strong conclusions related to causal relationships which could not be done based on the available research material. Therefore the internal validity of the research can be considered to be fulfilled at least to some extent.

It cannot be said that the results of this research which were conducted from individual cases could be generalized into broader theory. Therefore external validity cannot be fully assured. However, it can be stated that the study was basically replicated five times and that these five individual cases seemed to follow the hypothesis of this research. Therefore at least some level of generalization can be made but a much larger and wider research would be needed in order that the results could be really tested and possibly generalized into a broader theory.

The used research approach and methodologies are described in this research in detail. On top of that also the data collection method, the themes and questions used in the semi-structured interviews, and the research material are defined in this research. Therefore it can be assumed that another researcher could perform the same research and reach very similar results. The only aspect which needs to be noted regarding the reliability is the fact that the case companies insisted to be anonymous. Therefore another researcher could not actually go to interview the exactly same case companies. In order to compensate the anonymousness of the case companies the case descriptions were written in a very much detail. All in all, the reliability of this research can be considered to be fulfilled.

8. CONCLUSIONS

The main objective of this research was to prove the hypothesis that the use of comprehensive customer information and advanced analytical applications creates insight which can be used to further increase sales. The research starts from a literature review which includes such concepts as business analytics, business intelligence, business intelligence architecture, customer data, and customer analytics. On the contrary, the empirical research is based on five cases which were used to analyze the correlation between the business intelligence architecture maturity, ´the insightfulness of customer analytics, and increased sales.

This chapter includes the main results of the study and the conducted implications to practice and to theory. Additionally, the chapter contains the assessment of the whole study and ideas for further research.

8.1. The main results

All the five case companies of this research were evaluated based on the business intelligence architecture maturity model which was composed as a part of the theoretical part of this research. The business intelligence architecture maturities of the case companies varied a lot especially regarding specific subcategories of the business intelligence architecture maturity. The first case company was on the lowest maturity stage (2,8) while the fifth case company was on the highest maturity stage (4,2). Especially interesting were the differences in subcategories related to the comprehensiveness of data (the data sourcing maturity) and to the sophistication of analytical applications (the analytical technologies maturity). The first case company had the lowest maturity stage (2 - child) in the data sourcing category while the fifth case company had the highest maturity stage (5 - sage). On top of this, the first case company had also the lowest maturity stage (2 - child) in the analytical technologies category while the third and fifth case companies both were on the highest maturity stage (5 - sage).

Furthermore, also the amounts of different application areas of customer analytics in the case companies were evaluated. This was done based on the different application areas of customer analytics which were introduced also as a part of the theoretical part of this research. These included application areas related to customer identification, customer attraction, customer retention, customer development, and so on. On top of this, all the additional application areas of customer analytics which were raised up by the interviewees were included to the evaluation. Altogether 17 different application areas of customer analytics were practiced in the case companies. However, there were large differences between the case companies. The first case company practiced five different application areas of customer analytics while the fifth case company practiced 15 different application areas of customer analytics.

When the business intelligence architecture maturities of the case companies were compared to the amount of application areas of customer analytics it seemed that there is a correlation between the two. If the case company had a low business intelligence architecture maturity, it also practiced small amount of application areas of customer analytics and vice versa. Furthermore, when the subcategories of business intelligence architecture maturity were compared separately to the amount of application areas of customer analytics it was noticed that especially the data sourcing maturity and the analytical technologies maturity followed a similar pattern as the amount of application areas of customer analytics. Additionally, the average maturity of the data sourcing and the analytical technologies categories seemed to follow almost identical pattern as the amount of application areas of customer analytics. Therefore there are strong signs that at least the maturities of the data sourcing and the analytical technologies categories are related to the amount of different application areas of customer analytics a company is able to practice.

In order to understand the correlation between the business intelligence architecture maturity and the insightfulness of customer analytics, the level of customer analytics' sophistication was evaluated for each case company. This was done by considering the comprehensiveness and the integration into business processes of the different application areas of customer analytics. This evaluation showed that there are significant differences in the level of customer analytics' sophistication between the case companies. The first case company received the lowest (5) level of customer analytics' sophistication while the fifth case company received the highest (37) level of customer analytics' sophistication.

It also looked that the business intelligence architecture maturity is connected to the level of customer analytics' sophistication. This seemed obvious when the two were compared to each other. Both followed the similar pattern meaning that when the business intelligence architecture maturity was low also the level of customer analytics' sophistication was low and vice versa. As with the amount of application areas of customer analytics also with the level of customer analytics' sophistication it seemed that especially the data sourcing and the analytical technologies categories were connected to the level of customer analytics' sophistication. Additionally, when the data sourcing category maturity, the analytical technologies category maturity, and their average maturity was compared to the level of customer analytics' sophistication for each case company they followed a similar pattern. Especially the average maturity of the data sourcing and the analytical technologies categories followed the same pattern as the level of customer analytics' sophistication. Therefore there are also strong signs that at least the maturities of the data sourcing and the analytical technologies categories are related to the level of customer analytics' sophistication a company is able to practice. Thus, the most important result of the study was that the use of comprehensive customer information and advanced analytical applications seems to increase the insightfulness of company's customer analytics.

Furthermore, it was also validated that as described in the theoretical part of this research most of the application areas of customer analytics were ultimately related to increasing sales or sustaining current sales. Therefore if the insightfulness of customer analytics is connected to the maturity of the data sourcing and the analytical technologies categories of the business intelligence architecture, so is the increase or retention of sales which can be achieved with the support of customer analytics. Thus, the use of comprehensive customer information and advanced analytical applications seems to create insight which can be used to further increase sales.

8.2. Implications to practice and theory

The most important thing this research gives to its reader is the insight about the correlation between the business intelligence architecture and customer analytics. As certain categories of the business intelligence architecture seem to be important for the business value which can be achieved with the help of customer analytics it is beneficial to keep especially those categories in mind when developing customer analytics. It is however important to keep in mind that the other categories of the business intelligence architecture cannot be forgotten either. Additionally, there exist also multiple other factors which affect customer analytics besides technology related factors. These include such elements as the skills of the company's analysts, sponsorship, culture, funding, development processes, organizational objectives, and so on. Thus, if the results of this research are applied to practice, it is really important to remember that this study only considered the affecting factors related to the business intelligence architecture even though in practice there exist also multiple other affecting factors.

The research also provides one possible framework for evaluating company's business intelligence architecture maturity. The business intelligence architecture maturity model which was composed in the theoretical part of this study contains the main characteristics of a company in a specific stage of the business intelligence architecture maturity. By using these characteristics companies can evaluate their own maturity stage and possibly use the results of the maturity assessment to guide in high level when selecting development areas related to the business intelligence architecture and business analytics. However, the research does not provide the actual means how a certain company can move from a specific maturity stage to the following stage. Instead, it only provides the main characteristics of companies in different stages.

Additionally, both the results of the research and the theoretical part of this research can be used to acquire an understanding from customer analytics and some of the different possible application areas. For example the application areas which the case companies practice could be especially interesting when companies consider customer analytics. Most of the application areas of customer analytics, which are introduced in this research, are applicable to a wide variety of different companies and industries. Therefore companies can use the results of the research when they are considering should they start to practice customer analytics or when they evaluate their current stage of customer analytics in order to find future development areas.

Furthermore, the results of the research can also be used into some extent as one of the justifications to practice customer analytics. This is due to the fact that the research also considered the actual business value of customer analytics. This included especially increasing sales and sustaining current sales. Of course the research does not prove that customer analytics provides measurable business value but at least the subjective opinions of the interviewees seem to indicate such business value. Additionally, also many authors who are presented in the theoretical part of this study seem to see such business value achievable with the help of customer analytics.

The theoretical yield of this research is focused on to the different application areas of customer analytics and to the business intelligence architecture maturity model. The different application areas of customer analytics which were practiced in the case companies support the other researches related to customer analytics. The same application areas which were identified in other researches were also indentified in this research.

Additionally, a new maturity model related to business intelligence architecture was composed in this research from three widely used maturity models related to business intelligence and business analytics. This model was also shortly tested with the case companies as the business intelligence architecture maturities of the case companies were determined by using the maturity model. In this research the maturity model correctly found differences between the case companies and allowed the comparison between the business intelligence architecture maturity and the level of customer analytics' sophistication.

Furthermore, the most important theoretical yield of this research is the identified correlation between business intelligence architecture maturity and the insightfulness of customer analytics. Especially comprehensive data and advanced analytical applications seem to be closely connected to the company's ability to practice different application areas of customer analytics and to the level of sophistication of these practiced application areas of customer analytics.

8.3. Assessment of the study and further research

The research related to business intelligence and business analytics is challenging as the terminology especially related to business analytics is not fully stabilized. Some researchers write using business intelligence as an umbrella term which covers everything from the data management all the way to reporting and analytics. On the contrary, other researchers make clear distinction between for example business intelligence and business analytics. The researcher however considers that the theoretical part of this research was able to define what these and related terms mean in this specific research. All in all, the theoretical part of this research includes most of the relevant subjects when the scope of the research is considered. It can be however criticized that the research takes too narrow technology oriented approach when considering customer analytics. The researcher however has acknowledged this and considered also the other major factors affecting to the causal relationships between different conditions.

The research material was versatile and therefore from the researcher's opinion it was possible to identify much more interesting causal relationships between the business intelligence architecture and customer analytics than if the research material would have included for example only one case company. Of course in this situation the researcher would have been able to study one case in much more detail. It was however assessed that multiple cases would more appropriate as the goal was to prove the research hypothesis. Also the fact that case companies were from different industries was considered to be a good thing as it allowed the researcher to get a wider picture from the research topic. Even though some of the application areas of customer analytics may be more easily applied into certain industries customer analytics overall is not related to only specific industries.

This research has focused on to the identification of the correlation between business intelligence architecture maturity, customer analytics, and sales. This was however done based on five individual case companies. Therefore a much larger and wider research is needed in order that the results of this study could be really tested and generalized into a broader theory.

Furthermore, as stated this research focused on to the technology oriented aspects of customer analytics. There however exists also multiple other factors which affect customer analytics such as the skills of the company's analysts, sponsorship, culture, funding, development processes, and organizational objectives. It would be interesting and beneficial to compare the results of this study to other studies which would have focused on to the other factors affecting customer analytics. Therefore there is a need for further study on these other factors.

Additionally, further study would be needed related to the specific application areas of customer analytics. In this research customer analytics was studied widely as a whole but it would be also interesting to study a specific application area of customer analytics. This would allow for instance a deeper understanding about the causal relationships between business intelligence architecture or the other mentioned affecting factors and the insightfulness of that specific application area. Such an application area could be for example brand and product management which is based on to advanced text mining applications which extract customer insight from both external online media and internal CRM data.

BIBLIOGRAPHY

Arantola, H. 2006. Customer insight: Uusi va?ine liiketoiminnan kehitta?miseen. Helsinki, WSOYpro, WS Bookwell Oy. 158 p.

Boyer, K. K. & Hult, G. T. M. 2005. Customer Behavior in an Online Ordering Application: A Decision Scoring Model. Decision Sciences. Vol. 36, No. 4, pp. 569-598.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. & Wirth, R. 2000. CRISP-DM 1.0: Step-by-step data mining guide. CRISP-DM consortium. 78 p.

Chan, C. C. H. 2008. Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer. Expert Systems with Applications. Vol. 34, pp. 2754-2762.

Chen, Y., Tang, K., Shen, R. & Hu, Y. 2005. Market basket analysis in a multiple store environment. Decision Support Systems. Vol. 40, pp. 339-354.

Cho, Y. H., Kim, J. K. & Kim, S. H. 2002. A personalized recommender system based on web usage mining and decision tree induction. Expert Systems with Applications. Vol. 23, pp. 329–342.

Choo, C.W. 2002. Information management for the intelligent organization: The Art ofS-canning the Environment. 3rd edition. Information Today Inc., Medford, New Jersey, US. 325 p.

Cunningham, M. J. 2002. Customer Relationship Management. United Kingdom, Oxford, Capstone Publishing. 124 p.

Davenport, T. H. & Harris, J. G. 2007. Competing on analytics: the new science of winning. Boston, Mass, Harvard Business School Press. 215 p.

Davenport, T. H. & Harris, J. G. 2007b. The dark side of customer analytics. Harvard business review. May 2007, 9 p.

Davenport, T. H., Harris, J. G. & Morison, R. 2010. Analytics at work: smarter decisions, better results. Boston, MA, Harvard Business Press. 214 p.

Davenport, T. H. & Prusak, L. 1997. Information Ecology: Mastering the Information and Knowledge Environment. New York, Oxford University Press. 255 p.

Eckerson, W. W. 2007a. Interpreting Benchmark Scores Using TDWI's Maturity Model. TDWI Benchmark guide, TDWI Research, July 2007. 14 p.

Eckerson, W. W. 2007b. Predictive Analytics – Extending the Value of Your Data Warehousing Investment. TDWI Best Practices Report, First Quarter 2007. 32 p.

Eckerson, W. W. 2010. Four Ways to Build a Data Warehouse. Business Intelligence Best Practices. http://www.bi-bestpractices.com/view-articles/4770. Date of access: 16.11.2010.

Fayyad, U. 2003. Optimizing Customer Insight. Intelligent Enterprise Magazine. May 13, 2003, 6 p.

Gartner. 2009. Gartner's Business Intelligence, Analytics and Performance Management Framework. http://www.gartner.com/DisplayDocument?id= 1209327. Date of access: 18.10.2010.

Gonzalez-Wertz, C. 2009. The path forward: New models for customer-focused leadership - The 2009 Global CRM Leaders Study. USA, IBM Institute for Business Value, IBM Global Business Services, Executive Report. 16 p.

Hackathorn, R. 2004. The BI Watch: Real-Time to Real-Value. DM Review, January 2004. 4 p.

Hannula, M. & Pirttimäki, V. 2005. A Cube of Business Information. Journal of Competitive Intelligence and Management. Vol. 3, No. 1, pp. 34-40.

Hauser, W. J. 2007. Marketing analytics: the evolution of marketing research in the twenty-first century. Direct Marketing: An International Journal, Vol. 1, No. 1, pp. 38-54.

Hsieh, N. 2004. An integrated data mining and behavioral scoring model for analyzing bank customers. Expert Systems with Applications. Vol. 27, pp. 623–633.

Hwang, H., Jung, T. & Suh, E. 2004. An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. Expert Systems with Applications. Vol. 26, pp. 181-188.

Inmon, W., Strauss, D. & Neushloss, G. 2008. DW2.0: The Architecture for the Next Generation of Data Warehousing. USA, Burlington, Morgan Kaufmann Publishers. 371 p.

Jaspersoft 2010. The Analytics Landscape An overview of analytic technologies and how Jaspersoft‘s products and ecosystem address those areas. White Paper. 30 p.

Järvinen, P. & Järvinen, A. 2004. Tutkimustyön metodeista. Opinpajan kirja, Tampere. 211 p.

Kamakura, W. A., Wedel, M., Rosa, F. & Mazzon, J. A. 2003. Cross-selling through database marketing: a mixed data factor analyzer for data augmentation and prediction. Intern. J. of Research in Marketing. Vol. 20, pp. 45–65.

Kasanen, E., Lukka, K. & Siitonen, A. 1993. The Constructive Approach in Management Accounting. Journal of Management Accounting Research. Vol. 5, pp. 243-264.

LaValle, S. 2009. Breaking away with business analytics and optimization: New intelligence meets enterprise operations. USA, IBM Institute for Business Value, IBM Global Business Services, Executive Report. 13 p.

LaValle, S. 2009b. Business analytics and optimization for the intelligent enterprise. USA, IBM Institute for Business Value, IBM Global Business Services, Executive Report. 16 p.

LaValle, S., Hopkins, M., Lesser, E., Shockley, R. & Kruschwitz, N. 2010. Analytics: The new path to value - How the smartest organizations are embedding analytics to transform insights into action. USA, IBM Institute for Business Value & MIT Sloan Management Review, Executive Report. 21 p.

Lenzen, R. 2004. Customer Analytics: It's all about behaviour. DM Review, June 2004. 4 p.

Li, S., Sun, B. & Wilcox, R. T. 2005. Cross-Selling Sequentially Ordered Products: An Application to Consumer Banking Services. Journal of Marketing Research. Vol. 42, Issue 2, pp. 233-239.

Liautaud, B. & Hammond, M. 2001. e-Business intelligence: turning information into knowledge into profit. US. McGraw-Hill, 306 p.

Neilimo, K. & Näsi, J. 1980. Nomoteettinen tutkimusote ja suomalainen yrityksen taloustiede. Tutkimus positivismin soveltamisesta. Yrityksen taloustieteen yksityisoikeuden laitoksen julkaisuja. Sarja A 2: Tutkielmia ja raportteja 12. Tampere, Tampereen yliopisto. 82 p.

Ngai, E. W. T., Xiu, L. & Chau, D. C. K. 2009. Application of data mining techniques in customer relationship management: A literature review and classification. Expert Systems with Applications. Vol. 36, pp. 2592-2602.

Näsi, J. 1980. Ajatuksia käsiteanalyysista ja sen käytöstä yrityksen taloustieteessä. Yrityksen taloustieteen ja yksityisoikeuden laitoksen julkaisuja. Sarja A 2: Tutkielmia ja raportteja 11. Tampere, Tampereen yliopisto. 37 p.

Olkkonen, T. 1994. Johdatus teollisuustalouden tutkimustyöhön. Espoo, Teknillinen korkeakoulu, teollisuustalous ja työpsykologia. Raportti 152. TKK Offset, 114 p.

Pirttimäki, V. 2007. Business Intelligence as a Managerial Tool in Large Finnish Companies. Doctoral Dissertation. Tampere, Tampere University of Technology. 129 p.

Raisinghani, M. 2004. Business Intelligence in the Digital Economy – Opportunities, Limitations and Risks. USA, Idea Group Publishing. 289 p.

Rygielski, C., Wang, J. & Yen, D. C. 2002. Data mining techniques for customer relationship management. Technology in Society. Vol. 24, pp. 483–502.

Schroeck, M. J. 2001. Customer Analytics Making the Difference in CRM: Customer Analytics Amplify the Value of Integrated CRM Solutions. DM Review, September 2001. 6 p.

Spangler, S., Chen, Y., Proctor, L., Lelescu, A., Behal, A. & He, B. 2009. COBRA–mining web for COrporate Brand and Reputation Analysis. Web Intelligence and Agent Systems: An Internaltional Journal. Vol. 7, pp. 243-254.

Stone, M., Bond, A. & Foss, B. 2004. Consumer Insight: How to Use Data and Market Research to Get Closer to Your Customer. United Kingdom, Kogan Page Limited. 288 p.

TDWI. 2009. TDWI's Business Intelligence Maturity Model. http://download.101com.com/pub/tdwi/Files/TDWI_BIMaturity0609_lettersize.pdf. Date of access: 05.11.2010.

TDWI. 2010. What Works in Data Integration, The Data Warehousing Institute, Volume 29, May 2010, 43 p.

Thierauf, R. J. 2001. Effective Business Intelligence System. West Port, Quorum Books. 370 p.

Turban, E., Aronson, J. E. & Liang, T. 2005 (1993). Decision Support Systems and Intelligent Systems. 7th edition. Upper Saddle River, New Jersey, US. Pearson Education, Inc. 936 p.

Turban, E., Sharda, R. & Delen, D. 2011 (1993). Decision Support and Business Intelligent Systems. 9th edition. Upper Saddle River, New Jersey, US. Pearson Education, Inc. 696 p.

Turban, E., Sharda, R., Aronson, J. E. & King, D. 2008. Business Intelligence – A Managerial Approach. Upper Saddle River, New Jersey, US. Pearson Education, Inc. 225 p.

Voudouris, C., Owusu, G., Dorne, R., Lesaint, D. 2008. Service Chain Management - Technology Innovation for the Service Business. Springer. 308 p.

Wang, H. & Hong, W. 2006. Managing customer profitability in a competitive market by continuous data mining. Industrial Marketing Management. Vol. 35, pp. 715-723.

Woo, J. Y., Bae, S. M. & Park, S. C. 2005. Visualization method for customer targeting using customer map. Expert Systems with Applications, Vol. 28, pp. 763–772.

Yin, R. K. 2003. Case Study Research: Design and Methods. 3rd Edition. Thousand Oaks, SAGE Publications. 181 p.

Zaman, M. 2005. Predictive Analytics: the Future of Business Intelligence. Technology Evaluation Centers. 7 p.

APPENDICES (3 pieces)

APPENDIX 1: Detailed business intelligence architecture maturity

Stage / Category	1 - Infant	2 - Child	3 - Teenager	4 - Adult	5 - Sage
(a) - Data sourcing	-individual unintegrated internal data sources -missing data	-few integrated internal data sources -departmental data sets	-multiple integrated internal data sources -comprehensive business unit level data	-large amount of integrated internal data sources -also external data sources -consolidated enterprise level data	-large amount of integrated internal data sources -large amount of external data sources -fully comprehensive data -aware of the full scope of information available
(b) - Data quality & integration	-manual data extracts -non-trustworthy data -no data management	-First ETL-tools -recent data missing from data marts -small data quality improvements	-improved data integration and master data management capabilities -somewhat timely information -somewhat trustworthy data	-highly capable data integration -timely available data -data quality related issues mainly fixed	-sophisticated ETL-processes and master data management -fully trustworthy and timely data
(c) - Data warehouse architecture & scope	-manually maintained spreadmarts -individual scope -multiple different data definitions	-non-integrated data marts -departmental scope -inconsistent data definitions	-non-integrated data warehouses -business unit / division scope -somewhat consistent data definitions	-unified and centralized data warehouse architecture -enterprise scope	-data services available for any applications -inter-enterprise scope -consistent data definitions throughout the organization
(d) - Analytical technologies	-none	-isolated analytic efforts -stand-alone analytical tools -OLAP	-first initiatives to use sophisticated analytical technologies such as data mining tools which are able to identify patterns	-wide use of analytics -sophisticated analytical technologies which are able to make predictions	-highly sophisticated analytical technologies with functionalities such as predictive dashboards and business rules management - integrated into business processes
(e) - Business intelligence technologies	-operational reports and extracts produced by operational systems -spreadsheets	-isolated business intelligence efforts -stand-alone reporting tools -ad hoc queries & standard reports	-proliferation of business intelligence tools such as dashboards and scorecards	-centralized business intelligence tools which are used through a common portal	-business intelligence services available for any applications -business intelligence services available also for partners and customers

APPENDIX 2: The basis of the semi-structured interview

The start of the interview:

1. What is your current job and what kind of background you have?
    - career, title & tasks

2. Overall experience related to business analytics and to the use of analytic tools and applications?

Business intelligence architecture:

3. What kind of data warehousing environment your organization has?
    - data management
    - transformation tools and applications
    - data warehousing

4. What kind of analytical and presentation applications your organization has?
    - data mining, text mining, predictive analytics, simulation, rule engines
    - data access & reporting, information delivery

Customer insight & sales:

5. What kind of customer insight your analytical applications are able to provide?
    - in sales
    - in marketing
    - in service

6. How the provided customer insight helps your organization to create further sales?
    - through new identified up-selling and cross-selling opportunities
    - through targeted and automated marketing
    - through increased customer loyalty
    - etc.

Final questions:

7. What in your opinion should be the next step(s) when improving your organization's analytical capability from the technology perspective?
8. What kind of application areas of customer analytics you see as most promising in the future?
9. Would you like to add something regarding the theme?

APPENDIX 3: Comparison between the business intelligence architecture maturities of the case companies