MASTER

The business value of big data
a framework proposal for the financial services industry

de Vries, H.A.

Award date:
2013

Link to publication
The Business Value of Big Data

A Framework Proposal for the Financial Services Industry

by

Anke de Vries

Student identity number 0753144

in partial fulfilment of the requirements for the degree of

Master of Science

in Operations Management and Logistics

Supervisors:

1st: Dr. C.M. Chituc, TU/e, IS F. Pommée, Capgemini, FS

2nd: Dr. ir. H. Eshuis, TU/e, IS W. Ten Harmsen van der Beek – Hamer, Capgemini, FS
TUE. School of Industrial Engineering

Series Master Theses Operations Management and Logistics

Subject headings: Big Data, Business Information Management, Business Intelligence, Business Value, Financial Services Industry, Performance Management
Abstract

Nowadays organizations rely almost solely on Information Technology (IT). For example paying for products and getting insurance is all done with the use of IT. Organizations store enormous amounts of data that are saved for many years. Because of the fierce competition and economic downturn it is time for organizations to optimize the use of all the data that is already stored. One of the new technologies that can provide a solution is Big Data. However, it only is profitable to invest in Big Data when it provides value to the organization. The objective of this research is to develop a framework that determines how Big Data can add value to an organization in the financial services industry. A document study and interviews are performed to construct a conceptual framework for the business value of Big Data in the financial services industry. This conceptual framework is validated via a questionnaire among employees of Capgemini with knowledge of Big Data and the financial services industry. Last, a final value framework is proposed.

Big Data can be defined by the use of the three V’s volume, variety, and velocity. Big Data characterizes itself by the immense amounts of data, the use of structured and unstructured data, and the speed of the data. The fourth and fifth V, veracity and value, are not generally accepted, though this research work shows that Big Data indeed can provide value. Benefits of using Big Data are the more complete and (near) real time insights on an operational, financial, and business level, accurate response to customers’ needs, consistency of decision making, better risk management, and gaining competitive advantage. Challenges that should be taken into account are both technological as well as managerial. From a technological side it is a challenge to integrate information from different silos, to implement the complete technology, and to integrate it with the existing systems. From a managerial side it is a challenge to create a business culture in which Big Data is valued, and to have fully skilled employees. More general challenges that should be considered are the issue of privacy and security, and compliance with rules and regulations.

These benefits and challenges correspond with the benefits and challenges identified for the predecessor of Big Data which is Business Intelligence. Both technologies provide the overall package to transform data into information that is useful for the organization. However, whereas Business Intelligence can only transform structured data into information, Big Data can also transform unstructured data into information. As the definition already indicates, the data that the Big Data technology uses is bigger, faster, and has more variety than data that is used for Business Intelligence. Therefore it is more difficult to gain the benefits and deal with the challenges. The difference between Business Intelligence and Big Data is also shown in the types of systems that are used. Business Intelligence retrieves information from RDBMS while Big Data uses systems that do not need predefined queries. Contemplating the benefits and challenges of Big Data, the investment ultimately has to be profitable for the organization. In general business value can be described as everything that adds value to an organization. In this
research work business value is defined as everything that adds value to an organization in the financial services industry by using Big Data. Multiple approaches such as metrics and benchmarks could be used to determine the business value, but this research work focuses on frameworks because this is the most comprehensive method to assess the business value of Big Data in the financial services industry and literature lacks information about this subject. It is chosen to focus on one industry because the business processes and therefore the creation of value are different for every industry. The financial services industry is selected for this research work because of the rapidity in this industry which makes it interesting to discover if Big Data can have an influence.

To determine the business value of Big Data it is not yet possible to use quantitative frameworks such as the Performance Prism, because Big Data is not mature enough. Moreover, frameworks that include infrastructures and actors are also not possible to use because these are not completely established yet. It is also not possible to use Business Intelligence value frameworks as a reference because research on Business Intelligence is more mature. Therefore the framework the Benefit Logic Model developed by Capgemini is used to assess the business value. This framework provides a cause and effect diagram that shows how value can be created by implementing certain solutions. The solutions that should be implemented to create business value from investing in Big Data in the financial services industry are installing the hardware and software that Big Data requires, performing analytics and visualization techniques that are possible with Big Data, making a project team to create culture, methods and provide training, and attracting new employees that fit best with a Big Data environment. Together with drivers such as customized deals, cross selling, and fraud detection this will lead to an increase in revenue and decrease in costs. However, the preconditions privacy, security, and compliance to rules and regulations have to be met to create business value.

Comparing the Benefit Logic Model with other frameworks such as the Balanced Scorecard, they all incorporate business strategy and business operations. Differences between the Benefit Logic Model and other frameworks are that other frameworks provide objectives and measurements, and that different actors are involved. It is possible to use the Big Data Value Framework for the Financial Services Industry as a reference to develop a Big Data value framework for other industries, because often only processes differ between the industries. Additionally, the strength of the factors and their relationships will also differ across industries.

Limitations of this research work are that Big Data is a relatively new subject which makes it difficult to find solid articles that provide an overview of the new technology. Moreover, the validation of the conceptual framework is performed using a very small sample of the population which makes it difficult to generalize the findings. Last, another limitation is the lack of identifying drivers at the cost side of the framework. To deal with these limitations, recommendations and directions for further research are proposed at the end of this research work.
Preface

Today Information Technology is very important to organizations. The rise of the use of Information Technology (or in short IT) is the result of a growing internet organization culture in which business partners have to be fully connected all the time. Moreover, traditional business models are changed to more decentralized, flexible models that increase the use of Information Technology. Also globalization is a factor that influences the rise of using IT. That is why the area of IT is changing all the time. One of the newest developments in this area is the use of IT to make an organization more intelligent by providing information that can help managers to make better decisions. This development started with the introduction of Business Intelligence and now is grown to a platform in which technology and business come together, called Business Information Management. Big Data is the newest development in this field. It has similarities with Business Intelligence, but in some aspects it is different. Because Big Data is a new development, it is a subject in which organizations are currently interested. However, for an organization to invest in such a technology it should be profitable.

This research work determines how investing in Big Data can add value to an organization in the financial services industry.

This research is performed at the IT company Capgemini together with the University of Technology Eindhoven. At Capgemini it was possible to gain knowledge about Big Data by talking to different employees, sending the questionnaire, and using all documentation. Therefore I would like to thank everybody that helped me with the research and the organization as a whole for supporting this research work by providing the opportunities to perform this research. I would like to make a special thanks to the persons I interviewed because they helped me to develop the framework. It should be noted that not only employees of Capgemini allowed me to interview them, but also employees of other financial organizations. I would also like to make a special thanks to them. Last, I would like to thank my supervisors at Capgemini, F. Pommée and W. Ten Harmsen van der Beek – Hamer, for guiding me through this research, providing feedback, and putting in all the efforts to support me throughout the length of this process. I would also like to thank my supervisors at the University of Technology Eindhoven, C.M. Chituc and H. Eshuis, for the same reasons.
# Table of contents

Abstract ..................................................................................................................................................... i
Preface ........................................................................................................................................................ iii
Table of figures ........................................................................................................................................... vi
Table of tables .......................................................................................................................................... vii

1. Introduction ........................................................................................................................................ 1

2. Scope and Methodology ....................................................................................................................... 2
   2.1 Scope and research questions ......................................................................................................... 2
       2.1.1 Capgemini ................................................................................................................................ 2
       2.1.2 Financial Services ................................................................................................................ 2
       2.1.3 Research questions .............................................................................................................. 3

2.2 Research method ............................................................................................................................... 4
   2.2.1 Step 1: Document study .......................................................................................................... 5
   2.2.2 Step 2: Interviews .................................................................................................................. 6
   2.2.3 Step 3: Conceptual framework ............................................................................................... 6
   2.2.4 Iterations .................................................................................................................................. 6
   2.2.5 Step 4: Case study .................................................................................................................. 7
   2.2.6 Step 5: Evaluation ................................................................................................................... 8
   2.2.7 Outline ..................................................................................................................................... 8

3. Review on Business Intelligence, Big Data, and Value Assessment ................................................... 9
   3.1 The relation between Business Intelligence and Big Data ........................................................... 9
       3.1.1 Business Intelligence ........................................................................................................... 9
       3.1.2 Big Data .......................................................................................................................... 12
       3.1.3 Business Information Management ................................................................................... 17
   3.2 Business value of Business Intelligence ....................................................................................... 17
       3.2.1 IT business value .............................................................................................................. 18
       3.2.2 Corporate Performance Management .............................................................................. 18
       3.2.3 (Business) value of Business Intelligence ........................................................................ 20

3.3 Concluding remarks .......................................................................................................................... 22

   4.1 (Business) value of Big Data ....................................................................................................... 24
Table of figures

Figure 1: Process flow of research work.............................................................................................................5
Figure 2: Components of a data warehouse.........................................................................................................11
Figure 3: The three V’s of Big Data......................................................................................................................12
Figure 4: Steps in development of Benefit Logic Model........................................................................................27
Figure 5: Conceptual framework for applying Big Data in the financial services industry..............................31
Figure 6: Credit approval process........................................................................................................................34
Figure 7: Value from Big Data for the credit approval process.............................................................................35
Figure 8: Differences between conceptual framework and Big Data Value Framework for the Financial Services Industry........................................................................................................................................42
Figure 9: Big Data Value Framework for the Financial Services Industry.............................................................42
Figure 10: Big Data Value framework for the Financial Services Industry divided into four perspectives of Balanced Scorecard.............................................................................................................44
Figure 11: Development of Big Data........................................................................................................................67
Figure 12: First draft of conceptual framework....................................................................................................77
Figure 13: Second draft of conceptual framework................................................................................................78
Figure 14: Example bad P-P plot (1).......................................................................................................................86
Figure 15: Example bad P-P plot (2).......................................................................................................................86
Figure 16: Example good P-P plot..........................................................................................................................86
Figure 17: Example good histogram........................................................................................................................86
Figure 18: Example bad histogram (1)....................................................................................................................87
Figure 19: Example bad histogram (2)....................................................................................................................87
Table of tables

Table 1: Characterization of experts .........................................................29
Table 2: Solutions and drivers derived from interviews with experts ..............29
Table 3: Cronbach’s Alpha ........................................................................39
Table 4: Strong relationships (weight > 0.75) ............................................40
Table 5: Relationships that are considered neutral or less (weight < 0.7) .........41
Table 6: Descriptive statistics dataset ..........................................................79
Table 7: Descriptives gender ......................................................................81
Table 8: Descriptives age ...........................................................................81
Table 9: Descriptives nationality .................................................................82
Table 10: Descriptives education .................................................................82
Table 11: Descriptives current sector ..........................................................82
Table 12: Descriptives current company .....................................................83
Table 13: Descriptives job status .................................................................83
Table 14: Descriptives number of years worked in current organization .......83
Table 15: Cronbach’s Alpha for separate variables ......................................84
Table 16: Part of correlation matrix (1) .......................................................87
Table 17: Part of correlation matrix (2) .......................................................88
Table 18: Part of correlation matrix (3) .......................................................89
Table 19: Part of correlation matrix (4) .......................................................89
Table 20: Part of correlation matrix (5) .......................................................90
Table 21: Results of questionnaire .............................................................90
Table 22: Weights of lines after adaptation .................................................92
1. Introduction

Nowadays every organization uses Information Technology. To compete in a globally-integrated economy, organizations need a comprehensive understanding of markets, customers, products, regulations, competitors, suppliers, employees and more (IBM, 2013). Information Technology is a subject that has been researched for years and is still an important topic for research today. One of the developments in IT is Big Data which involves transforming data into knowledge that ultimately contribute to the value of an organization. Characteristics of Big Data are the enormous volume, variety, and velocity of data (Capgemini, 2012; Database Trends and Applications, 2012; De Vries, 2013b; IBM, 2013; Lopez, 2012; Won, 2013). Google alone already processes 24 petabytes of data per day (Davenport, Barth, & Bean, 2012). Big Data is used in a variety of organizations, for example, the police in Los Angeles uses Big Data to predict where crime will most likely happen in the next 12 hours. Moreover, during the last campaign of president Obama, Big Data was used to select the floating voters and only those persons were personally approached. Also organizations in the financial services industry use Big Data, for example credit card companies have live use cases of Big Data. On the other hand, there are also limitations, for example the company Equens developed a new business model but stopped to introduce this business model due to public opinion. Though, only using the technology without a clear business vision does not provide the value organizations aim at (Bean, 2013). Strategic possibilities and advantages are only possible to achieve when IT and business work together. Business – IT alignment is “applying Information Technology (IT) in an appropriate and timely way, in harmony with business strategies, goals, and needs” (Luftman, 2000). Business – IT alignment consists of strategies concerning decisions referring to competitive choices and choices that apply to the capabilities of an organization to execute the choices (Henderson & Venkatraman, 1999).

A solid business-IT alignment can lead to better business performance as discussed by Gregor, Hart, and Martin (2007). To create a strong relationship between business and IT, investments in IT should contribute to the business goals and performance. Therefore, investments in IT should create business value. Business value could shortly be described as the organizational performance impacts that are caused by the use of IT. This research work will provide a framework that indicates how business value can be created by investing in Big Data in the financial services industry.

This research work is divided into six chapters. The next chapter includes the scope and methodology used to conduct this research work. The third chapter encompasses a review of Business Intelligence, Big Data, their relationship, and value assessment. The fourth chapter contains a framework proposal for the business value of Big Data in the financial services industry. The fifth chapter explains how the conceptual framework is validated including a comparison with other frameworks. The final chapter includes the conclusion, limitations, and directions for further research.
2. Scope and Methodology

In this chapter the methodology that is used for this research work is explained. First, the context in which this research is performed is described. Additionally, the research questions guiding this research work are provided. Last, all steps taken during this research are clarified.

2.1 Scope and research questions

Every research work is scoped to make it possible to be more in-depth about the subject that is researched. This study focuses on the business value of Big Data in the business area of financial services. An empirical investigation is performed at the organization Capgemini.

2.1.1 Capgemini

Capgemini is an IT organization which is focused on consulting, technology, outsourcing and local professional services. The organization has its origin in 1962 as the French IT company CAP Europe. The organization Capgemini is started in 1996 after multiple mergers and acquisitions. It is an international organization with more than 100,000 employees and the headquarters is located in Paris. The area in which Capgemini is active is Europe, Asia, North- and South – America. Capgemini offers specific solutions across six sectors from which Financial Services is one of them. The financial services industry is an important sector for Capgemini together with the public sector and the retail industry. Different solutions are provided for amongst others risk management and compliance, global delivery network, and financial services partnerships (Capgemini, Financial Services, 2012).

The vision of Capgemini is to create value for customers by developing the right solutions to support the goals of the customer’s company which are made by a set of experts who work closely together. Moreover, Capgemini strives to achieve concrete business results for the customer in order to be more flexible and better resistant to competition. To accomplish this, collaboration is very important. Therefore, experts of the company work closely together with the employees of the customer’s company (Capgemini, 2013a).

2.1.2 Financial Services

Financial services is a generic term for all economic services provided by the finance industry. The financial services industry encompasses amongst others banks, foreign exchange services, investment services, insurance companies, credit card companies, and consumer finance companies. The areas Capgemini focuses on are compliance and risk management, insurance, payments, retail banking, wealth management, agile, and optimizing business applications.

Credit card companies are one of the main investors in Big Data (Nasar & Bomers, 2012). However, all parts of the industry are adapting to the trends in financial services. In the last decade many new regulations are introduced in the banking and insurance area. Additionally, competition has grown which forces companies to change the attention from simple services for the private sector to more profitable activities as public offerings and investments for the business sector. To comply with these trends, a different kind of employee is necessary. Integrity
is one of the most important competencies next to expertise in the financial area because the work is more specialized and complex. Moreover, a commercial attitude is important because of the rising competition. The increasing competition leads to a further internationalization of the financial markets. Furthermore, the customer has the power to decide which company will be chosen. This pushes organizations to add elements to their financial service to provide a total solution for financial problems instead of just a single service. Additionally costs have to be reduced because of fierce competition and the economic downturn. All these trends push organizations to be more customer oriented. This also has an influence on the IT that is used. The IT systems have to be more flexible to meet the requirements of the customers (Thielen, 2013).

2.1.3 Research questions
Different articles show different explanations of the IT feature Big Data. Defining the business value of Big Data is a subject that has not been researched extensively yet. Analytical tools coming from Big Data can make complex insights easier to understand and make this information immediately ready to use at every point in the organization and at every skill level (LaValle et al., 2011). Big Data can lead to more effective fact-based decision making and optimized business processes. By creating these new types of analytical tools, an organization can be helped to sell more products and provide better service (Capgemini, 2012). In general, Big Data can create business value. However, investing in such a new technology has to be considered thoroughly to ensure that the investment will be profitable. Therefore this research work will contribute to existing literature by providing a framework that helps to determine the business value of Big Data in the financial services industry. To achieve this objective, the following research questions have been defined, which will guide the research work to be pursued. Research question one, two, and three will set the stage to answer research question four which relates to the objective of this research.

1. What is Big Data?
   a) What are the definitions of Big Data?
   b) What are the positive aspects of Big Data?
   c) What are the challenges of Big Data?
   d) How is Big Data incorporated in an organization?

2. What is the relationship between Business Intelligence and Big Data?
   a) What are the similarities between Business Intelligence and Big Data?
   b) What are the differences between Business Intelligence and Big Data?

3. What are the most relevant approaches to assess the business value of Big Data in an IT context?
   a) What is business value?
   b) What are possible approaches to determine business value in general?
4. How can the business value of Big Data be determined in the area of financial services?
   a) How can the business value of Big Data be determined?
   b) What are similarities with other frameworks?
   c) What are differences with other frameworks?

2.2 Research method

A research methodology is used to have a structured way of answering the research questions stated in the previous section. The methodology is guided by the research work of Hevner, March, Park, & Ram (2004) who describe a framework for understanding, executing, and evaluating Information Systems (IS) research. IS research operates on the crossroad of people, organizations, and technology, by combining behavioral and design science. Behavioral science describes research “through the development and justification of theories that explain or predict phenomena related to the identified business need” (Hevner et al., 2004, p. 79). On the contrary, design science “addresses research through the building and evaluation of artifacts designed to meet the identified business need” (Hevner et al., 2004, p. 79). An IT artifact can be a construct, model, method, or instantiation that enable researchers to understand the problems that come across in developing and successfully implementing information systems within organizations (March & Smith, 1995). So, the goal of behavioral science research is the truth, while the goal of design science research is utility. Though, truth and utility are inseparable, because truth informs design and utility informs theory. To improve the utility of the artifact or identify weaknesses in theory, evaluation activities are performed to refine and reassess. Hevner et al. (2004) identify three components of behavioral and design science research that interact with each other: the environment, IS research, and knowledge base. The environment defines the business needs. IS research contains the research activities which in case of design science imply building and evaluating artifacts that are designed to meet the business needs. It includes a build step in which the artifact is constructed and an evaluation step in which the artifact is validated. The knowledge base is the foundation that provides applicable knowledge for the build step of IS research and methodologies that provide knowledge for the evaluation step.

These three components are used as a base for this research work. First the environment is established. In the case of Capgemini, the need of the design of a framework is posed that provides the business value of Big Data in the financial services industry. Additionally the knowledge base is used as a foundation to build the artifact which is the conceptual framework. A document study and interviews are performed (step 1 and 2) to build this framework (step 3). This process is repeated for several times to optimize the conceptual framework. The knowledge base is also used to provide methodologies for the evaluation step. The methodology that is chosen is the case study which is performed via a questionnaire (step 4 and 5). The steps that are taken are shown in figure 1 and explained in the paragraphs below.
2.2.1 Step 1: Document study

The first step performed during this research is a document study. Different sources of literature are reviewed and discussed to examine the topic of this research. The document study contributes to the conceptual framework that is the basis for research in practice by providing elements that have an influence on the business value of Big Data. Therefore a solid review of scholarly literature is performed. Subjects where literature is searched for are, amongst others, Business Intelligence, Business Information Management, Big Data, financial services industry, performance measurement, business value of Business Intelligence, and business value of Big Data. Different kind of databases could be used to perform a literature study. According to Sekaran (2003) there are bibliographic databases, abstract databases, and full-text databases. Bibliographic databases only present bibliographic citations such as the name of the author, year of publication, and page numbers. Abstract databases additionally show an abstract or summary of the article. Full-text databases also display the full text of the article. For this research work only full-text databases are used. Various retrieval systems like ABI/Inform, ScienceDirect, Web of Knowledge and Emerald are used to find high quality full-text articles about the subjects of interest. Also Google Scholar is used to find in-depth articles about the subjects. Many articles were found, but only the articles which are focused on the main subjects were used. The magazines from which articles are used are, amongst others, MIS Quarterly, Information System Research, European Journal of IS, Information Systems Management, and Decision Support Systems. Additional to the articles found in the different magazines, also books and conference proceedings are used that are related to the different topics of this research. The articles, books, and conference proceedings that were found are used as a starting point for the research.

Additionally, content-related documents provided by Capgemini and other organizations are used to gather additional information about the topics. During this part of the research information is collected that is used as input for step 2 and step 3.
2.2.2 Step 2: Interviews

The second step that is taken during this research is performing interviews. There are three types of interviews: structured, semi-structured, and unstructured interviews. A structured interview is conducted when it is known what information is needed. A set of predetermined questions is used to collect information from the respondent (Sekaran, 2003). A semi-structured interview refers to the context where questions are generally phrased and it is possible to vary in the sequence of the questions. Moreover, the interviewer has the freedom to ask additional questions that seem relevant but that are not predetermined (Bryman & Bell, 2007). The unstructured interview has the characteristic that the interviewer only has a list of topics that can be discussed and no predetermined questions at all. The sequence of the questions is totally dependent on the interview itself (Bryman & Bell, 2007).

During this research, eight semi-structured interviews are performed with experts in the area of financial services and Big Data from Capgemini and customers of this company. At one interview two persons were questioned at the same time. The purpose of these interviews is to acquire information about their vision on Big Data, possible approaches to assess the business value of Big Data, the benefits an organization would gain by Big Data, and feedback on the how to develop a conceptual framework. The advantage of using semi-structured interviews is that this type of interview gives the companies’ perspective on the issue as well as it could confirm insights that come from the performed document study. A set of open questions were conducted to introduce topics, but it remained exploratory of nature to obtain new insights and to allow discussion. The questions are shown in appendix A. Moreover, the output of the interviews is used to develop and improve the conceptual framework.

2.2.3 Step 3: Conceptual framework

Based on the information obtained during step 1 and step 2 it is possible to construct a conceptual framework that indicates how business value can be created from investing in Big Data. The artifact is developed for the business need to discover the value of Big Data in the financial services industry.

2.2.4 Iterations

After step 3 an iteration takes place which indicates that step 1, 2, and 3 are (re)built and evaluated in order to improve the quality of the conceptual framework. In total two iterations are performed during this research. The first draft of the conceptual framework includes all the factors that are discovered in the document study and the first five interviews. This resulted in an enormous framework that contains a lot of factors that are not specifically related to Big Data or the financial services industry. Based on discussions with the supervisors it was decided to perform step 1, 2, and 3 again. This time the document study was focused on finding factors specifically for the financial services industry. The last three interviews were also more aimed at finding factors for this industry. Together this resulted in a second draft of the conceptual framework. Also this framework was proposed during a meeting with the supervisors and it was decided that the framework is still too large. The focus on the financial services industry is well
incorporated, however there are still some factors included that are not directly related to Big Data. Therefore a second iteration took place to delete the factors that are not directly related to investing in Big Data but that are more of a general IT nature. To decide which factors should be deleted, again a study was performed to evaluate the already found documents and interviews to delete all the factors that were not mentioned in a Big Data setting. This resulted in the third, more compact version of the conceptual framework. During a meeting with the supervisors it was decided that this framework represents the business value of Big Data in the financial services industry and could be used as input for the questionnaire.

2.2.5 Step 4: Case study

Hevner et al. (2004) describe that evaluation is a crucial component of the research process. Five design evaluation methods are discussed: observation, analytical, experimental, testing, and descriptive. For this research work, the observational evaluation method is most applicable, because the evaluation occurs within Capgemini and the artifact will be analyzed within its environment. The observational evaluation method consist of a case study or field study, which differ in the number of projects in which the artifact is evaluated. For this research work, it is chosen to perform a case study because the evaluation occurs at one project, and not at multiple projects.

A case study “involves in-depth, contextual analyses of similar situations in other organizations, where the nature and definition of the problem happen to be the same as experienced in the current situation” (Sekaran, 2003, p.35). In this case study, various experts of Capgemini with knowledge of Big Data and the financial services industry are questioned about the same problem. These experts are different from the experts that are interviewed. Case studies that are qualitative in nature are useful for understanding certain phenomena and acquiring theories for empirical testing (Sekaran, 2003). The case study is aimed at increasing the credibility and applicability in practice of the proposed framework by finding the empirical support needed. Current research work on the objective of this research is not yet available, therefore finding empirical support is even more important.

The research method that is used to perform the case study analysis is the questionnaire. A questionnaire is “a preformulated written set of questions to which respondents record their answers, usually within rather closely defined alternatives” (Sekaran, 2003, p. 236). The questionnaire is used to validate the conceptual framework constructed at step 3. To minimize the time that is needed to complete the questionnaire it is rather short and to the point. The questionnaire is electronic and distributed via e-mail. At the end of the questionnaire it is possible for respondents to provide comments.

The questionnaire will be more attractive by writing a good introduction, organizing the questions, giving instructions and guidance, and aligning the questions to make sure the respondent can complete the task expending least time and effort (Sekaran, 2003). Also a conclusion of the questionnaire will make it more attractive. These characteristics are taken into
account while making the questionnaire. Also the type of response categories is taken into account while designing the questionnaire. Open-ended questions are avoided as much as possible. Although open-ended questions allow the respondent to answer in any way, more time has to be spent to think about all possible answers that could be given by the respondent. Closed questions are questions in which respondents are asked to make choices among a set of alternatives, which can be yes/no questions, multiple choice questions, and scale-answers. The advantage is that respondents only have to consider the options already given by the researcher. Moreover, closed questions are easier to analyze. A disadvantage is that the researcher has to be sure that all possible answers are provided. Therefore it has to be ensured that the alternative answers are mutually exclusive and collectively exhaustive (Sekaran, 2003). The questionnaire can be requested via the contact details in appendix B.

2.2.6 Step 5: Evaluation
The last step of this research is the evaluation. During this step information from the case study is analyzed and evaluated to validate and improve the conceptual framework about the business value of Big Data in the financial services industry. Moreover, the framework is compared to other frameworks and models that are used for trends preceding Big Data to indicate what is similar and what is different.

2.2.7 Outline
The steps identified in the paragraphs above are used as a guideline for this research work. Information derived from step 1, the document study, is used in chapter three and four to provide information about the different subjects. Also information gathered by the interviews (step 2) is used in chapter three and four. Designing and improving the conceptual framework (step 3) is shown in chapter four. Information from the last steps, step 4 and step 5, are both provided in chapter five which explains how the case study was performed and shows the final results.
3. Review on Business Intelligence, Big Data, and Value Assessment

This chapter provides a review on Business Intelligence and Big Data and the relation between these technologies. Business Intelligence can be considered as the predecessor of Big Data and is therefore chosen as the starting point for this research. Additionally, Business Information Management, performance management, business value and value frameworks for Business Intelligence are discussed. At the end of this chapter the concluding remarks are provided. This chapter provides a general overview, for detailed information see appendix C.

3.1 The relation between Business Intelligence and Big Data

Although Business Intelligence and Big Data are different, they both start with the same: data. Data can be described as “raw, unprocessed streams of facts” (Benson & Davis, 2008, p. 8). Organizations collect all kinds of data such as facts about their daily operations, products, or customers. Raw data is a relative term in the sense that data can be processed in a number of stages in which processed data can be considered as raw data for the next stage. When data is transformed into a meaningful form that can be used by a person or computer it is called information. Data and information are distinguished by the business context in which they are used (Benson & Davis, 2008). Other authors (English, 1999; Sen, 2001; Curtis & Cobham, 2002; Laudon & Laudon, 2012) agree with this distinction between data and information and emphasize that data is part of and needed to get information. Davenport (1997) already mentioned that it is possible to construct a continuum of data, information and knowledge which in the end can lead to different definitions. Besides data and information, Davenport (1997) also introduces the term knowledge as the superlative of information.

Business Intelligence, Business Information Management and Big Data represent the overall technology and business package that transforms data into information and knowledge to ultimately create value for the organization. The next paragraphs provide information on the different subjects and their relationship.

3.1.1 Business Intelligence

Definition

The first term that comes across in history about using data intelligently is Business Intelligence. In literature there are different, semi-identical definitions used for the term Business Intelligence. Business Intelligence as an umbrella term is introduced in 1989 by H. Dresner of the Gartner Group who described it as ‘all the technologies that help business make decisions based on fact’; ‘Using fact rather than intuition was the key to intelligence’ (Nylund, 1999).

According to Olszak and Ziemba (2003) Business Intelligence systems consist of different modules. The systems contain tools to extract and transfer data, data warehouses to store data, analytic tools (OLAP) to access, analyze and model the data, tools to report the analysis, and a presentation layer that provides information to users in a comfortable way. Negash (2004) defines Business Intelligence as “combining data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to
planners and decision makers” (p. 178). Jourdan, Rainer, and Marshall (2008) support this by presenting Business Intelligence as “both a process and a product. The process is composed of methods that organizations use to develop useful information, or intelligence, that can help organizations survive and thrive in the global economy. The product is information that will allow organizations to predict the behavior of their competitors, suppliers, customers, technologies, acquisitions, markets, products and services, and the general business environment with a degree of certainty”(p. 121).

Different definitions are given, but the main point that is followed during this research work is that Business Intelligence contains the mixture of raw data and analytical tools to present useful information that supports organizations in making decisions. This raw data can be structured or unstructured. Structured data is data that fits neatly into relational or flat files, while unstructured, also called semi-structured, data does not fit decently into relational or flat files. Business Intelligence uses structured data for the analysis of data and delivers the information to the decision maker (Negash, 2004). Other terms are also used for Business Intelligence, such as Competitive Intelligence, Corporate Intelligence, Competitive Information, or Commercial Intelligence. These terms emphasize the competitive advantage that can be taken by the use of intelligence. Although the terms Competitive Intelligence and Business Intelligence are used interchangeably, there is a difference. McGonagle & Vella (2002) describe the difference by discussing that Business Intelligence is only oriented to the internal processes, while Competitive Intelligence is also oriented to the external processes. According to Negash (2004), Competitive Intelligence is a specialized branch of Business Intelligence.

**Benefits and challenges**

The goal of Business Intelligence is to handle large amounts of information and to identify and develop new opportunities. This can provide a competitive advantage in the market and long-term stability (Rud, 2009). So, Business Intelligence integrates analysis of data with decision-making tools to provide the right information, on the right time, in the right location, to the right persons throughout the organization (Ghazanfari, Jafari, & Rouhani, 2011). In short, it supports organizations to gain business advantage from data.

Benefits of using Business Intelligence can be financial related such as cost and time savings. Other benefits are identifying threats and opportunities, understanding the vulnerability of the organization, decrease reaction time, and out-think competition (Thomas Jr., 2001). Willen (2002) shows that Business Intelligence supports in strategic and operational decision making by using it for corporate performance management, optimizing customer relations, monitoring business activity, traditional decision support, and management reporting of business intelligence. In the field of financial services, Business Intelligence systems can be used to improve efficiency and visibility in the back-office operations, because all levels of management can see how problems are related (De Voe & Neal, 2005). Dave (2009) divides the advantages of using Business Intelligence into three categories: revenues, customer relationship, and sales and marketing. Business Intelligence distinguishes the products and services that drive revenues and
it ranks customers based on their profitability. Moreover, it will detect customer relationship problems early and takes action to resolve them. Last, Business Intelligence can be used to rank successful promotions based on product and market segments, and to discover what is in the sales pipeline. Business Intelligence also provides challenges, for example the databases that are used have to be current, timely, and accurate to benefit from it (Thomas Jr., 2001). Moreover, another challenge is that there will be costs involved to operate complete Business Intelligence, such as hardware costs, software costs, implementation costs, and personnel costs. A comprehensive budget should be available to cover these costs from Business Intelligence (Negash, 2004).

Technology

Over the years data warehouses became the most important technology for Business Intelligence. The transformation of data into information happens by integrating data from different sources done in a data warehouse (Inmon, 2013). A data warehouse is an independent business database system that is filled with data that is extracted from different sources in order to improve the decision-making process within an organization (Curtis & Cobham, 2002). The data warehouse consolidates and standardizes the data to make it applicable for management analysis and decision-making (Laudon & Laudon, 2012). A data warehouse consists of several components which are shown in figure 2.

![Figure 2: Components of a data warehouse (Adapted from: Laudon & Laudon (2012))](image)

The Extract and Transform component extracts current and historical data from different internal and external operational systems in the organization. The Information Directory component provides information about the data that are currently in the Data Warehouse. Last, the Data Access and Analysis component gives data access and options for analysis. It is often referred to as the local warehouse in which aggregated data from the Data Warehouse is stored (Laudon & Laudon, 2012).
The process that is followed with Business Intelligence contains four steps. First, data is gathered from different systems and placed into a data warehouse. Second, the gathered data is transformed to make sure that the data is uniform and can be compared. Next, the collected data is analyzed to transform it into information that is usable for the management. Last, the found information is presented via a dashboard, for example to improve the decision making process (Laudon & Laudon, 2012).

3.1.2 Big Data
According to Arnold (2012) and Capgemini (2013b) organizations use Business Intelligence applications that provide information to the core of their businesses. However, these applications only represent a small part of the data that can be analyzed. Business Intelligence tools are good in analyzing structured data while organizations also store many unstructured data that can provide information.

Definition
Quite recently a new term is introduced in Business Information Management that allows systems to transform unstructured data into information: Big Data. One generally accepted description of Big Data is based on the three V’s: volume, velocity, and variety (Capgemini, 2012; Database Trends and Applications, 2012; De Vries, 2013b; IBM, 2013; Lopez, 2012; Won, 2013). The volume refers to the enormous amounts of data that are gathered on a daily basis, for example from Twitter. Velocity reflects the speed level at which data is streamed. Variety stands for the numerous sources of data that is streamed into, out of, and through organizations like transactions, social media and websites. Next to that, this streaming happens in a big range of different formats from traditional data to message documents, videos, and audio files. The three V’s are also shown in figure 3.

These three V’s separate Big Data from other trends (De Vries, 2013b). The data volumes are higher than what is processed before, the volumes are larger than what is handled by traditional database technology, external data is added to the existing internal data, various data is combined to analyze, and possibilities are created to gather information near real-time (Capgemini, 2013e).
Different sources claim that there is a fourth V: veracity (IBM, 2013). Veracity represents the level of trust that the information on which decisions are made is correct. Some sources also indicate that there might be a fifth V which represents the value that could be created by using Big Data to enable enhanced decision making, to get insight discovery, and to optimize processes (De Vries, 2013b). However, the fourth and fifth V are not generally accepted.

A different description of Big Data is shown in Capgemini (2013e) that defines Big Data by three elements which are the data itself, the process for dealing with data, and the holistic view it can enable. Although there are different definitions, Big Data can be described as “massive amounts of stored content (structured or unstructured) that can be easily analyzed in real time (a reasonable amount of time to get a useful answer)” (Arnold, 2012, p. 32). According to Boyd and Crawford (2012), Capgemini (2013e), and De Vries (2013b), the term Big Data only refers to the quantity of data, forgetting the growing velocity and variety that separates it from Business Intelligence. That is why Boyd and Crawford (2012) define Big Data as an interplay of the technology which maximizes computational power to gather and compare large data sets, the analysis which identifies patterns to make economic, social, technical or legal claims, and the belief that large data sets offer intelligence with the aura of truth, objectivity and accuracy.

Benefits and challenges
Big Data advances the core technologies needed to collect, store, preserve, analyze, manage, and share enormous amounts of data (Lazar, 2012). Data can be collected across business units, from partners, and customers while a flexible infrastructure integrate all this information effectively. Analytics finally create sense in all of this information (Brown, Chui, & Manyika, 2011). These predictive analyses are important to create value from Big Data. Using new technologies on existing data can provide additional dimensions of that data. Moreover, new types of analytics can help organizations to provide an improved overall customer experience, implement analytics in real time, and yield competitive intelligence about other offers (Capgemini, 2013e). Moreover, Big Data provides organizations with opportunities that were not feasible before because of the lacking technology or high expenses (Bakshi, Bhambhri, Vesset, & Zedlewski, 2013; Lopez, 2012).

Big Data gets increasing attention in the financial services industry because accurate, consistent management of both financial data and customer information is essential to be successful in this industry (Nasar & Bomers, 2012). It is striking that the volume of the data is often not the problem for organizations in the financial services industry. Organizations are used to work with large amounts of data, but the growing velocity and variety are more difficult to cope with. The variety of data can have an impact on organization’s risk measurements and its trading and investment performance. The difficulty lies in integrating the information from different sources by combining structured and unstructured data (Brown, 2012; Bean, 2013; Capgemini, 2013b; Capgemini, 2013e). The velocity of data is most important because it represents to what extent an organization can filter relevant information from the other two V’s as fast as possible to
develop new products, to optimize processes, and to make better decisions (Capgemini, 2013e; De Vries, 2013b).

According to Lopez (2012), using Big Data can translate into operational, financial, and business gains, including one of the ultimate benefits which is that an organization can have quicker access to cleaner, more relevant data to drive insights and optimize decision making. Capgemini (2013e) adds to this that the use of Big Data allows organizations to analyze a much broader set of data about aspects of the organization as an integrated whole. Combining data from inside and outside the organization provides a more holistic view on different aspects in the organization such as customers, supply chains, products, and processes. Capgemini (2013b) indicate these benefits as a more complete understanding of market conditions and evolving business trends, more accurate responses to customer needs, consistency of decision making, better risk management, and a competitive advantage. LaValle, Lesser, Shockley, Hopkins, & Kruschwitz (2011) and Database Trends and Applications (2012) add to this that the real benefit of Big Data is that it provides an opportunity to businesses by enabling users to run analytics and determine and predict market shifts, customer preferences, and product innovations. These benefits are also applicable in the financial services industry, however there is also another benefit from using Big Data analytics concerning regulatory compliance. Using Big Data to analyze the current regulatory environment can help organizations comply with today’s increasingly complex and everchanging regulatory requirements (Nasar & Bomers, 2012). However, this is difficult in the financial services industry because the rules and regulations often change faster than the IT landscape (De Vries, 2013b). Capgemini (2013e) summarizes these benefits in three parts: improving interaction with the ecosystem that results in better targeted marketing, individual service offerings and customer retention, improving business processes to understand the data more in detail and better predict future activity, and risk mitigation to manage risk compliance which is especially important in the financial services industry.

To gain from these benefits Big Data can provide, organizations need a solid plan. Organizations need to identify their key business drivers, how to put the processes into place, and need to make sure that they have the organizational alignment and skills to make it happen. In short, to be successful an organization should have the right people with the right skills on the right place and the right organizational structure to use Big Data (Bean, 2013).

Besides the benefits Big Data can provide, there are also challenges to the new development. One of the challenges is to integrate all information from different data silos together without separating it again in different silos. As said in Capgemini (2013b) and in De Vries (2013b), the biggest obstruction to use Big Data effectively for decision-making is the organizational silos. Additionally, today’s relational database architectures need to be modernized to adapt to the increasing volume, variety, and velocity without removing the existing systems (Database Trends and Applications, 2012; DBTA, 2013a). Nasar and Bomers (2012) add that designing, developing, and implementing data management systems that live up to the access and speed
requirements is a highly complicated process. Only when people, processes, and technology are all integrated it is possible to make the Big Data solution successful (Capgemini, 2013b; Capgemini, 2013e; De Vries, 2013b; Nasar & Bomers, 2012). This is also shown by LaValle et al. (2011), who indicate that many organizations have more trouble with the managerial and cultural barriers rather than data and technology barriers. To use analytic-driven insights, there should be a close connection to the business strategy, it should be easy to understand for the end-users, and it should be embedded into the processes, which are all difficult to implement. Brown et al. (2011), Capgemini (2013b), and De Vries (2013b) agree with this and indicate that one of the major complications for successfully implementing Big Data is talent. Organizations need people with deep analytical skills and managers with a clear understanding of how Big Data can be applied. So, additional substantial investments need to be made in recruitment and training. Moreover, time and money should be invested in change management that, amongst others, provides training to all employees to deal with the new business environment. The culture of the company should be changed to a culture in which the use of Big Data is more valued and rewarded (Capgemini, 2013b; De Vries, 2013b). This is only possible when every part of the organization is involved in the process and collaborates with each other (Capgemini, 2013e). However, cases that could be easily solved should not become more complex than they initially were (De Vries, 2013b).

Boyd and Crawford (2012) indicate challenges that are more general of nature. One of these challenges is the issue of privacy. Because a lot of information is available by using Big Data, this does not mean that it is ethical to use it all (Boyd & Crawford, 2012; Capgemini, 2013e; De Vries, 2013b). According to De Vries (2013b) there are only a few specific rules and regulations about what is allowed concerning the privacy of individuals. Organizations often pay more attention to what a customer accepts as to what extent information can be used than what is allowed, because customers often accept less than what is prescribed by the law. This is also confirmed by Brown et al. (2011) who indicate that greater access to personal information that is often demanded by Big Data put pressure on privacy and security issues. Therefore organizations have to be transparent about the use of the data of the customer. Moreover, it should be clear to whom the data belongs (De Vries, 2013b).

**Technology**

One of the main differences between Business Intelligence described above and Big Data is the difference in systems that are used. Whereas Business Intelligence retrieves information from Relational Database Management Systems (RDBMS), Big Data tries to combine data from different management systems together to provide more accurate information. Using distributed file systems that are available with Big Data technology allows organizations to get quick access to information. These Big Data technologies work best in cooperation with the original enterprise data warehouses as used with Business Intelligence (Nasar & Bomers, 2012). Existing infrastructures should not be dispensed, but current capabilities should be integrated with the new requirements of Big Data. This is also agreed upon by Capgemini (2013e) that indicates that
Big Data is not only about technology but more about data management that combines traditional techniques with newer ones. The four steps that need to be taken to implement a Big Data solution are acquisition, marshalling, analysis, and action. The acquisition step includes collecting data from a variety of sources using extraction, transfer, and integration. The second step, marshalling, encompasses sorting and storing the data depending on the intended use. The analysis step contains finding insights and predictive modeling by combining information from the past with forward analytics. The last step, action, includes taking action on the insights to change business outcomes carried out by a human, a computer, or a combination of both. Especially this last step can be taken faster with the use of Big Data compared to the traditional approaches. The visualization of data becomes increasingly important for actions involving a human interaction. Because data sets are too large to make sense to a human, the visualization of data, including patterns and graphs, is used to base decisions on (Capgemini, 2013e).

The current RDBMS are not capable of dealing with the new data types (especially the unstructured data) and do not provide a convenient way of storing data that is necessary for Big Data. Collecting and storing both structured and unstructured data should be possible without a clear idea how an organization can use it (DBTA, 2013b). Conventional relational databases are often structured and work with a well-defined schema. Using these databases for analysis is difficult when data sets are constantly changing because they are designed to optimize repeated queries. A new technology that does not use these repeated queries is the Hadoop technology (Bakshi et al., 2013; Capgemini, 2013e). The technology that belongs to the core of the Big Data revolution is MapReduce which was developed by Google to distribute and execute the computations where the data is originally stored. Later, the technology became available as an open source version via Hadoop (Arnold, 2012). Hadoop divides information into different blocks and does not need a predefined schema. It is designed to deal with large data sets by distributing the data over many servers enabling reliable, scalable, and distributed computing on these clusters of data (Capgemini, 2013e). Hadoop looks for patterns in data and dealing with unpredictable data sets and therefore focuses on flexibility and experimentation. Once a file is imported and stored in the system it can be accessed when it is needed. However, Hadoop lacks in providing direct responses when many people query at the same time. Organizations are used to work with Business Intelligence tools which provide an instant response on predefined questions, but Hadoop cannot provide that instant response (Bakshi et al., 2013). However, using Hadoop means that the data engineering process can be skipped and therefore it is possible to answer critical business questions in a shorter amount of time (Bean, 2013). Additionally, Hadoop in itself does not create significant value. Hadoop provides the possibility to manage and process large amounts of data, but it should be complemented with other tools and platforms to optimize the data management. Hadoop only addresses the storage and analysis steps excluding the acquisition of data and actions that can be taken based on the analysis part (Capgemini, 2013e). Moreover, using Hadoop looks like a less expensive option than adding nodes to database appliances, however hardware is still needed to cope with the growing data
volumes and performance service level agreements which raise the capital and operational costs. Also the ease of use is a challenge with Hadoop. Certain actions require specific skills that are hard to find which makes the use of Hadoop more difficult for organizations (Lopez, 2012). Other programs that use similar technology are jStart, Vivisimo, Vetas, Aster Data Systems, Vertica, Netezza, and Greenplum (DBTA, 2013a; IBM, 2013).

3.1.3 Business Information Management

The term Business Information Management encompasses the management process of converging data into information to gain insights and foresights in business performance. Whereas Business Intelligence and Big Data mainly focus on the techniques to transform data into insights, Business Information Management mainly focuses on the management behind it. It combines people, processes, and technology to create value from intelligence. Business Information Management consists of four parts: Business Intelligence, Big Data, User Experience, and Data Management. These four parts should be combined to create business value. Business Intelligence and Big Data refer to the technology, Data Management includes the processes, and User Experience addresses the people part of Business Information Management. In literature, Business Intelligence and Business Information Management are often used interchangeably, though there is a difference in the focus of the terms (Negash, 2004; De Vries, 2013b).

According to Edgar (2005) Business Information Management provides a holistic understanding of management of information and information systems, the processes within a business, and the processes that transform data and information into knowledge. Business Information Management has the characteristic that it combines two perspectives: technological and business. The technological view includes the role of IT in the complete process; how IT can be used for gathering data, integrating data, analysis of data, and the distribution of information over the organization. This technological view refers more to the parts of Business Intelligence and Big Data. The business view includes the systematic process in which the core business strategy is the base for the business information needs that require data collection and analysis using knowledge and information. This knowledge and information then contribute to the business strategy. Capgemini (2013b) supports this and indicates that data can either help a manager in making a decision or it can automate decision-making. This business view refers more to the parts of Data Management and User Experience.

3.2 Business value of Business Intelligence

To discover the business value of Business Intelligence it first should be clear what business value means in an IT setting. Only when this is clear it is possible to determine the business value of Business Intelligence. Therefore next paragraph will provide information on IT business value. Additionally the business value of Business Intelligence is explained.
3.2.1 IT business value

As there are multiple ways to define Business Intelligence and Big Data, there are also multiple ways to define business value. In De Vries (2013b) business value reflects the performance that is visible and indicates what is actually performed and not what is said to be performed. In De Vries (2013b) business value is also described as everything that adds value to the business. Moreover it is described as the extent to which an organization is relevant to the customers in such a way that it is most efficient and effective on an emotional and rational level. More related to the financial services industry, in De Vries (2013b) business value is also described as having insights at the services that are provided and the acceptance of the customer.

A short definition comes from Masli, Richardson, Sanchez, and Smith (2011) who define business value as “the organizational performance impacts of IT” (p. 83). According to Hitt and Brynjolfsson (1996), and Devaraj and Kohli (2003) these organizational performance impacts of IT include productivity enhancement, cost reduction, inventory reduction, profitability improvement, competitive advantage, and other measures of performance. However, according to Melville, Kraemer, & Gurbaxani (2004) the term performance is used for both intermediate process-level measures and organizational measures. Additionally they discovered that two formulations of performance are used in literature, efficiency-based performance and effectiveness-based performance. Efficiency emphasizes an internal perspective using metrics such as cost reduction and productivity, and shows the actual output compared to the input. Effectiveness emphasizes the achievement of organizational objectives in relation to the external environment by comparing the expected and actual output. To be complete, Melville et al. (2004) define IT business value as “the organizational performance impacts of IT at both the intermediate process level and the organization wide level, and comprising both efficiency impacts and competitive impacts” (p. 287).

3.2.2 Corporate Performance Management

Determining business value is only possible when performance is measured. The part of an organization that deals with how business value can be created is Corporate Performance Management (CPM). CPM, also known as Enterprise Performance Management (EPM) or Business Performance Management (BPM), entails the continuous process in which methodologies, metrics, processes, and systems are used to manage and monitor performance by achieving goals that induce the ultimate goal of realizing the defined strategy (Buystendijk, 2005; Kemper, Rausch, & Baars, 2013). CPM functions as a communication device between the strategic level of management and the business operations on a tactical and operational level. Strategic objectives and performance targets are top-down communicated while reliable and consistent information is communicated bottom-up to feed the steering and control process (Van Wunnik, Heizenberg, Smits, & Pommée, 2006). The CPM process contains four steps. First, norms should be set for every measurement. Second, the actual measurements should be done. Third, the actual measurements have to be compared to the norms set during the first step to
discover potential deviation. The last step includes taking a certain action that is based on the conclusions from step 3 (Buylendijk, 2005).

Performance is often expressed by financial indicators, because the business decisions resulting from the performance outcomes are mostly based on financial measures. Moreover, financial indicators are more tangible than the non-financial ones which makes it easier to understand. However, focusing on financial indicators has some disadvantages. First of all, financial measures only indicate something about the organization itself, not about the competition. Additionally, financial indicators are based on historical data and do not provide information for the future. Last, using financial indicators is aimed at the shareholder, forgetting other stakeholders such as customers and employees.

These disadvantages are taken into account at different frameworks that could be used to manage performance. The research work of De Vries (2013a) shows many different performance frameworks. One of the frameworks that is widely known and used in practice is the Balanced Scorecard of Kaplan and Norton (1996; 2000). The Balanced Scorecard approach allows a comprehensive view of the nature of the IT investment. Moreover, it highlights how IT investments influence the processes and the overall organizational performance in one framework (Masli et al., 2011). According to Kaplan and Norton (1996) the traditional accounting models that could be used to measure the business value of IT do not incorporate the intangible assets and company capabilities such as high quality products and services, skilled employees, and satisfied customers that are critical for success. Therefore a different framework is developed that includes both the historical-cost financial accounting model and new long-range competitive capabilities, which is the Balanced Scorecard. It complements the financial measures of past performance with the measures of the drivers of future performance. The objectives and measures are derived from an organization’s vision and strategy. Organizational performance is viewed from four perspectives: financial, internal business process, customer, and learning and growth. All perspectives are divided in objectives, measures, targets, and initiatives. The objectives indicate what an organization wants to evaluate. The measures are used to realize the objectives. The targets give an indication of the current and expected value for measures of objectives that have to be realized. Last, the initiatives are assigned for each objective and the responsibility is determined (Kaplan & Norton, 1996).

Many organizations already use financial and non-financial measures for organizational performance, however the non-financial measures are often only used to improve locally at the front-line and at customer-facing operations. Financial measures are often used to only control short-term operations. The Balanced Scorecard emphasizes that both type of measures should be part of the information system for all employees at all levels of the organization. The measures are deduced from a top-down process driven by the vision and strategy of the organization and represent a balance between the external measures for shareholders and customers, and the internal measures. The scorecard incorporates a complex set of cause and effect relationships.
among the critical variables that describe the trajectory of the strategy. The financial results are driven by customer satisfaction which is driven by internal processes. Underneath these layers, there is the foundation of the learning and growth perspective (Kaplan & Norton, 2000). Moreover, it also provides a mix between outcome measures and performance drivers such as cycle time and defect rates. Furthermore, a feedback loop is included which makes sure that the objectives that are determined are questioned once in a while if they are still consistent with current observations and experiences (Kaplan & Norton, 1996).

3.2.3 (Business) value of Business Intelligence

Also for the Business Intelligence business value contains everything that adds value to the company. In literature there is no unanimity about how to measure the business value of Business Intelligence.

The paper of Lönnqvist and Pirttimäki (2006) indicates that it is difficult to get unanimity about how to measure the business value of Business Intelligence investments. First of all it should be clear what the purpose of the measurement is before it can be decided what and how to measure it. Different purposes for measuring performance are decision making, control, guidance, external communication, and education and learning (Simons, 2000). The main purposes for measuring the performance of Business Intelligence are to prove that the investment is worth it and to manage the process to ensure that users are satisfied and the process is efficient. Therefore it is necessary that it is clear how much it costs to apply Business Intelligence and what the benefits are (Lönnqvist & Pirttimäki, 2006).

To calculate the total costs, labor costs, information purchases, and other expenses should be included (Lönnqvist & Pirttimäki, 2006). The method Total Cost of Ownership proposed by Degraeve, Labro, & Roodhooft (2004) could be used to identify all relevant costs. However, according to Lönnqvist and Pirttimäki (2006), measuring the benefits is more difficult because benefits of the investments are often non-financial and even intangible, such as improved quality and timeliness of information. Additionally, it is even difficult to assess the financial benefits. The outcome of Business Intelligence is intelligence which is some kind of processed information. Determining the value of this information in financial terms is difficult. According to Lönnqvist and Pirttimäki (2006) it is better to measure the effectiveness. Four measures are identified: time savings, cost savings, cost avoidance, and revenue enhancement. The effectiveness can be measured by evaluating the contribution of Business Intelligence to a specific decision or action and how it benefits the organization. Moreover, it is possible to measure the effectiveness subjectively by evaluating the perceived customer satisfaction. One thing that should not be forgotten is that measuring the performance should also be done in the context of managing the process. The goal is to efficiently produce valuable intelligence for the specific needs of the users.

The paper of Williams and Williams (2003) also describes how different methods can be used to measure the performance. According to Williams and Williams (2003), often the business value
of an investment is measured as the net present value of the after-tax cash flows. Therefore investments in Business Intelligence should be subjected to an assessment of how the investment results in increased revenues and/or reduced costs. All benefits that are proposed such as agility, responsiveness, information sharing, and flexibility are meaningless when these attributes are not defined in operational terms and realized through business processes to influence cash flows. That is why the business value of Business Intelligence can be defined as how an organization uses Business Intelligence to improve management processes such as planning and monitoring to change the cash flows from a management side, and to improve operational processes such as sales campaign execution and purchasing to increase revenue and/or reduce costs from a business side. This means that to capture the business value, organizations should go beyond the technical implementation of the systems. Preconditions that should be met to capture the business value are strategic alignment, process engineering, change management, BI technical development, and BI project management. Only when these conditions are met it is possible to create business value for the specific environment. The business value does not only contain a traditional return on investment analysis which estimates the net present value of after-tax cash flows, but it should be taken from a broader perspective including opportunity analysis, readiness assessment, process engineering, and change analysis (Williams & Williams, 2003).

According to Lin, Tsai, Shiang, Kuo, and Tsai (2009), not only efficiency and effectiveness determine the business value of Business Intelligence, but it is also important that the systems can be accepted by the users, that it satisfies the needs of the users, and that the objectives that are composed are met after implementation. This can be measured by contents correctness, easiness of the operation, resilience of the format, integrity of the output, real-time nature, integration and safety of the system, and credibility of the output.

Although above measurements could be used to assess the business value of Business Intelligence, more comprehensive methods exist to have a better overview of all possible measures that influence the organizational performance. Literature proposes different balanced performance measurement frameworks that can be used to assess the business value of Business Intelligence, such as the Balanced Scorecard of Kaplan and Norton (1996; 2000). Lönnqvist and Pirttimäki (2006) show that the main principles are the same for every framework, dividing the principles over five stages. First performance measures are chosen based on an organization’s strategy. Next, the success factors are chosen from the different perspectives included in the framework. In the third stage measurement is limited to a number of critical success factors. After that the measurement system is designed in such a way that causal relationships between the success factors are shown. During the last stage the measurement system can be used as a tool to implement and communicate the initial strategy.

Although there are different performance frameworks that could be used to assess the business value, no framework is specifically developed for Business Intelligence. Additionally, different types of measurement are discussed, such as financial and non-financial measurements, but there is no agreement on the best way to assess organizational performance.
3.3 Concluding remarks
This chapter described the terms Business Intelligence, Business Information Management, and Big Data and their relation to each other. Moreover, value assessment in general and value assessment of Business Intelligence are discussed.

Business Intelligence contains the mixture of raw, structured data and analytical tools to present useful information that supports organizations in making decisions. Data warehouses are used to store and transform different types of data to analyze and transform it into usable information for the management. Advantages of using Business Intelligence are amongst others cost and time savings, identifying threats and opportunities, and gaining competitive advantage. Business Intelligence integrates analysis of data with decision-making tools to provide the right information, on the right time, in the right location, to the right persons throughout the organization. Challenges that occur when using Business Intelligence are that databases have to be accurate and investments have to be made into hardware, software, implementation, and personnel. Big Data can be seen as the successor of Business Intelligence and is generally described along the lines of the three V’s volume, velocity, and variety. Different sources claim that there are also a fourth and fifth V (veracity and value), but these are not generally accepted. Instead of using traditional RDBMS as used with Business Intelligence, Big Data uses systems that do not use repeated queries such as Hadoop. Though, the steps that are followed are the same. Using Big Data provides the same benefits as using Business Intelligence, however insights will be more complete because also unstructured data can be analyzed. Therefore it is easier to comply with rules and regulations, to better target the customers, and to identify risks. Because of the bigger data it is important to have the right people with the right skills on the right place and the right organizational structure. Improving the technical part of the organization (combining new and legacy systems, integrating information from different silos, visualization of information) is as important as improving the cultural part of the organization (skills, business culture). A challenge that is highlighted with the use of Big Data is the issue of privacy and security. Personal information should be secured at all times and used with care. Business Information Management is an overall term for the management process of converging data into information to gain insights and foresights in business performance. Whereas Business Intelligence and Big Data mainly focus on the techniques to transform data into insights, Business Information Management mainly focuses on the management behind it. Business Information Management consists of four parts: Business Intelligence, Big Data, User Experience, and Data Management. These four parts should be combined to create business value.

Business value can be described in many ways, but overall it can be seen as everything that improves the performance of an organization. This description of business value can be specified to a certain domain such as IT or the financial services industry. Questions about productivity, business profitability, and value for customers have to be answered to identify business value for a certain investment. Creating business value is only possible when performance is measured.
The part of an organization that deals with how business value can be created is Corporate Performance Management (CPM). It entails the continuous process in which methodologies, metrics, processes, and systems are used to manage and monitor performance by achieving goals that induce the ultimate goal of realizing the defined strategy. Multiple approaches could be used to construct business value. A comprehensive method that incorporates all possible measures of business value is the use of frameworks. Different performance management frameworks could be used such as the Performance Pyramid, Balanced Scorecard, and Performance Prism. Also the business value of Business Intelligence could be assessed, but there is no unanimity how to do that. Measuring the benefits is more difficult because benefits of the investment are often non-financial and even intangible. Therefore it is proposed to measure the effectiveness and efficiency of Business Intelligence. Also performance frameworks could be used to determine the business value of Business Intelligence, but there are no frameworks specified for this type of investment.

Business Intelligence can be considered as the predecessor of Big Data. Literature provides information about the business value of Business Intelligence, but lacks information about the business value of Big Data. Definitions, benefits, and challenges of Big Data are described, but no concrete process or framework is provided that shows how business value can be created from investing in Big Data. This research work will cover this gap by determining a value framework for investing in Big Data. It is chosen to focus on one industry, because the business processes and creation of value are different for every industry. Moreover, constructing a framework that is not specified for one industry will be too large to be useful. This framework will focus on the financial services industry. This industry characterizes itself by the enormous amounts of daily transactions and quickly changing rules and regulations. It is important that the IT in a financial organization is up to date to be able to deal with the rapidity of this industry. That is why it is interesting to discover if investing in Big Data can be profitable in this industry. Additionally, literature does not provide extensive information about the role of Big Data in the financial services industry. Therefore, developing a Big Data value framework specified for this industry will contribute to the limited existence of literature about Big Data in the financial services industry.

In this chapter the business value of Big Data is discussed. Although literature is not clear about gaining value from Big Data (the possible fifth V), this research work will clarify that investing in Big Data can provide value and how this value can be reached. First a review about the value of Big Data is presented together with the possible approaches and challenges to assess the value of Big Data in general. After that the conceptual framework is explained which is focused on the financial services industry. This conceptual framework is constructed via an iterative process of reevaluating the framework based on the information gathered by the document study and interviews.

4.1 (Business) value of Big Data

4.1.1 Introduction

The business value of Big Data again can be described as everything that adds value to the organization, but now with the use of Big Data. In De Vries (2013b) it is claimed that the business value of Big Data is not different from the business value of Business Intelligence which includes making the business more intelligent by making decisions based on good information. According to Bean (2013), the key metric that provides business value is the ability to facilitate time-to-answer using Big Data. Arnold (2012) adds to this that the value of Big Data is the possibility to analyze information that was impossible to analyze before. There are two main areas: where and how to store data, and how data can be accessed in real time. Usually both issues could be dealt with separately, but with Big Data it is possible to address them at once. The old data and systems will stay, but it will be easier to access the data. Organizations that use these new knowledge streams will have a competitive advantage.

Also Capgemini (2013d) agrees that Big Data can provide business value. The paper indicates three areas in which analytics using Big Data can have an influence: supply analytics, enterprise analytics, and customer analytics. Supply analytics include developing a better understanding of the spending in the organization, identifying opportunities to be more efficient, optimizing inventory, maximizing availability and service levels, and creating agile supply. Enterprise analytics contain a better understanding of the true drivers of costs and revenue, understanding the true profitability, combining financial and operational information to understand business outcomes, visualizing the business outcomes, making better (strategic) decisions, and creation of a pool of skilled resources. Last, customer analytics encompasses the understanding of the true customer profitability, better targeting and retaining of customers, and identifying the spending profiles of customers.

Additionally, DBTA (2013b) indicates that many organizations only focus on the technical issues and not on the other challenges that are important to create business value. These other challenges are defined by ‘the four T’s’ which are talent, trust, transformation, and tangible
results. Especially talent is critical, because organizations need the right blend of skills and knowledge such as technical capabilities, and deep industry expertise by combining data scientist skills with data analysis skills, and skills in improving algorithms. In De Vries (2013b) all these points are combined into five goals from which at least one needs to be reached to add value to the business. These goals are increasing revenue, decreasing costs, being more agile, being compliant to rules and regulations, and customer satisfaction. Big Data mainly influences the goals customer satisfaction and being more agile, though it probably has an indirect effect on increasing revenue and decreasing costs.

LaValle et al. (2011) created a five point methodology to successfully implement analytics driven management that is needed to create business value from Big Data. The first recommendation is to focus on the biggest and highest value opportunities. It is easier to justify a big effort when a potential big reward is in sight. However, only start the opportunities with strategic business direction, otherwise it will only provide costs. The second recommendation is to start with questions and not with the data. Instead of first gathering all the available data and contribute a lot of time to collecting, cleansing, and converting data that may not even be useful, organizations need to start with defining the insights and questions needed to meet the business objective and then collect data that can provide answers to this business objective. The third recommendation is to embed insights to drive actions and deliver value. The new methods make insights more understandable and actionable due to the data visualization, simulations, business process analytics, and advanced statistical techniques. The fourth recommendation is to keep existing capabilities and to add new ones, but do not detract any capability. The new and sophisticated tools should supplement the earlier ones, because it does not mean that the older tools are not useful anymore because some new tools are developed. The fifth and last recommendation includes that parts should be build, but the whole should be planned. An information agenda is needed to plan for the future. The information agenda is used to share and deliver trusted information across all applications and processes and aligns the IT and business goals by providing enterprise information plans that are financially justified.

There are three critical management needs that have to be met by all the five points: reduced time to value, increased likelihood of significant and enduring transformation, and focus on achievable steps. Value can already be created in the early phase of analytics and enables strategic and cultural change that lasts because it overcomes the most significant organizational hurdles. Moreover, focusing on achievable steps empowers organizations to focus their efforts and resources instead of implementing universal changes (LaValle et al., 2011).

4.1.2 Use of frameworks

To capture the business value of Big Data, frameworks can be used. The use of frameworks is a comprehensive method that provides an overview of all possible measures that influence organizational performance. Examples of performance measurement frameworks are the
Balanced Scorecard of Kaplan and Norton (1996; 2000), the Performance Pyramid of Lynch and Cross (1995), Kanji’s Business Scorecard of Kanji and Moura e Sá (2002), and the Performance Prism of Neely, Adams, and Kennerly (2002). These performance frameworks all have predefined measurements that are used to compare the organization’s results with predefined standards.

Frameworks that determine the business value of Business Intelligence could be used as a reference to assess the business value of Big Data, because Business Intelligence and Big Data both focus on making an organization more intelligent by the use of data. However, the development of Business Intelligence systems is more mature than the development of Big Data systems which makes it difficult to use frameworks of Business Intelligence. Because Big Data is often not used in practice yet, it is difficult to already implement quantitative performance frameworks at all. Therefore a more qualitative approach should be taken to identify the business value for Big Data. A possible approach for a more qualitative framework is the Business Model Canvas of Osterwalder and Pigneur (2009). The Business Model Canvas is a strategic management framework for documenting existing and new business models. Nine building blocks are included in the model: customer segments, value propositions, channels, customer relationships, revenue streams, key resources, key activities, key partnerships, and cost structure. These nine building blocks are categorized into infrastructure, offering, customers, and finances. The category infrastructure includes the key activities that are most important in executing the value proposition, key resources that are necessary to create value for the customer, and key partnerships that are necessary to optimize operations and reduce risks. The category offering encompasses the value proposition which is the collection of products and services that are offered to meet the needs of the customer. With the value proposition an organization distinguishes itself from its competitors. Value can be provided through various elements such as performance, design, newness, price, getting the job done, customization, status of brand, cost and risk reduction, usability, and accessibility. The third category, customers, contains customer segments that identify which customers an organization tries to serve, channels through which the value proposition can be delivered to the targeted customers, and customer relationships that identifies the type of relationship an organization wants to create with their customers. The last category, finances, include the cost structure that identifies the most important monetary consequences for different business models, and revenue streams that indicate how an organization gain income from each customer segment.

Another qualitative framework that could be used to identify the value for Big Data is the e3 Value Model of Gordijn (2004). The e3 Value Model is a methodology that helps to better understand the business model by involving all stakeholders, and performing an analysis and profitability assessment of the business model for all parties involved. Key for this methodology is the notion of economic value, and how actors create, change, and consume objects of economic value. It is used to bridge the gap between business and IT groups. The e3 Value Model is divided into three viewpoints: the global actor, the detailed actor, and the value activity.
The global actor viewpoint includes all actors involved, the objects of economic value, objects of value that are expected in return of an object of value delivered, objects that are offered or requested in combination, and phenomena that cause exchanges of objects between actors. The detailed actor viewpoint encompasses the partnerships between actors, the constellations of actors, and phenomena that influence this viewpoint. The value activity viewpoint contains the value-creating activities and their assignment to actors.

A last approach that could be used to assess the business value of Big Data is the Benefit Logic Model developed by Capgemini (De Vries, 2013b). Traditionally the Benefit Logic Model is used to define a cause-effect diagram that shows how solutions contribute to cash flow generation. Cash flow generation can be created by increasing revenue or decreasing costs. Breaking down these two sides into smaller parts will lead to improvement opportunities that can be reached by implementing the solutions. To create a Benefit Logic Model, a key design issue is necessary to scope the model. Additionally, the Benefit Logic Model is also used to identify the benefits of a certain investment. For example, the Benefit logic Model is used to show the opportunities for investing in Business Intelligence (De Vries, 2013b). Figure 4 shows how a Benefit Logic Model can be created.

![Figure 4: Steps in development of Benefit Logic Model](image)

### 4.1.3 Concluding remarks
Using Big Data can provide value by gaining more complete insights into different parts of an organization including customers, processes, and risk mitigation. Additionally it can provide a competitive advantage if it is used correctly. This is only possible when different challenges are met. First of all, organizations need to integrate the new Big Data applications with the old systems to use Big Data in an optimal way without making unnecessary costs. Moreover,
information has to be visualized to make sense to an organization. Another challenge that is more managerial than technical is that the business culture has to change to an open environment in which Big Data is valued. Additionally, employees need the right skills to deal with the enormous amounts of fast and diverse data. To overcome these challenges, complying with the rules and regulations, and privacy and security issues should be taken into account at all times.

To capture the value of Big Data into a model, different approaches could be used. A comprehensive method that incorporates all possible measures of business value is the use of frameworks. Different performance management frameworks could be used such as the Performance Pyramid, Balanced Scorecard, and Performance Prism. These frameworks could be adapted to the characteristics of Big Data by using Business Intelligence performance frameworks as a reference. However, Business Intelligence is more mature in its development than Big Data which makes it difficult to use these frameworks. Moreover, Big Data is not often used in practice yet which makes it difficult to measure the performance at all. Therefore qualitative frameworks are proposed that show how Big Data could provide value to an organization. A possible approach for a more qualitative framework is the Business Model Canvas which is a strategic management framework for documenting existing and new business models. Another framework that could be used is the e3 Value Model which is a methodology that helps to better understand the business model by involving all stakeholders, and performing an analysis and profitability assessment of the business model for all parties involved. A last approach that could be used to assess the business value of Big Data is the Benefit Logic Model which uses a cause-effect diagram that shows how solutions contribute to cash flow generation.

The Benefit Logic Model is chosen to use for the value framework of Big Data for the financial services industry because of the clear overview, the focus on value, and the structured approach of the model. Moreover, the development of Big Data is at its starting point which makes the Benefit Logic Model the best option to show how Big Data can be used to gain value. With the Benefit Logic Model it is possible to provide insight into the value of Big Data without other elements such as the infrastructure and actors. Additionally, as described in paragraph 3.3, it is impossible to create a Big Data value framework that is not specified for one industry. In this research work it is chosen to create a Big Data value framework for the financial services industry because of the rapidity of this industry. The Benefit Logic Model fits with this characterization, because the essence of the framework is not influenced by the rapidly changing rules and regulations.

4.2 Conceptual framework

As discussed in the paragraph before, the Benefit Logic Model will be the base for the conceptual framework. The key design issue is capturing the value of investing in Big Data for the financial services industry. The conceptual framework is based on information gathered by the document study and interviews with nine experts in the fields of financial services and Big Data. The characterizations of the experts are shown in table 1.
The information from the document study is used as input to define the questions for the semi-structured interviews. These questions that are used as a guideline for the interviews are shown in appendix A. The general information on Business Intelligence, Business Information Management, Big Data and Value Assessment gathered by the interviews with experts are combined with the document study in chapter 3. Specific information about factors that influence the value of Big Data is combined in table 2. This table shows the solutions that could be implemented right away and the drivers that lead to a positive cash flow. Moreover, it is shown how often a certain solution or driver is mentioned during the interviews.

Table 2: Solutions and drivers derived from interviews with experts

<table>
<thead>
<tr>
<th>Items derived from interviews</th>
<th>Solution</th>
<th>Driver</th>
<th>Precondition</th>
<th>Number of times mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer retention</td>
<td></td>
<td>X</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Customized offering</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Fraud/skimming/phishing detection</td>
<td></td>
<td>X</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Customer risk behavior /Risk profiling/Credit scoring</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Process insight (all processes in a bank such as borrowing, saving, cards etc.)</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Investing in change management</td>
<td>X</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Investing in technology of Big Data (hardware and software)</td>
<td>X</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Cross-selling</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Customer satisfaction</td>
<td></td>
<td>X</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>
Table 2: Solutions and drivers derived from interviews with experts (continued)

<table>
<thead>
<tr>
<th>Items derived from interviews</th>
<th>Solution</th>
<th>Driver</th>
<th>Precondition</th>
<th>Number of times mentioned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event-driven marketing</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Insights in behavior market/sentiment analysis</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Real time information delivery</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Increase of agility/velocity</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Improve company image</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Change in mindset of employees</td>
<td></td>
<td>X</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Attract new type of employees</td>
<td></td>
<td>X</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Discovering the total experience of customers (search-process-buy)</td>
<td>X</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Analytics and visualization of data</td>
<td>X</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Privacy</td>
<td></td>
<td></td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>Security</td>
<td></td>
<td></td>
<td>X</td>
<td>3</td>
</tr>
<tr>
<td>Compliance to rules and regulations</td>
<td></td>
<td></td>
<td>X</td>
<td>2</td>
</tr>
</tbody>
</table>

In the conceptual framework the solutions are grouped together on the left side and are linked to the added value by the drivers. Only the solutions and drivers that are mentioned more than once are used in the conceptual framework. The driver ‘event-driven marketing’ is combined with the driver ‘customized offering’ because it both aims at providing a personalized offer for a customer. The drivers ‘real time information delivery’, ‘increase of agility/velocity’, and ‘discovering the total experience of customers’ are linked to the driver ‘process insight’ because the separate drivers together indicate the same as the driver ‘process insight’. This driver includes (real-time) insights in the behavior of customers and the processes itself and delivers the information as fast as possible. The solution ‘change in mindset of employees’ is combined with the solution ‘investing in change management’ because it is part of change management.

With the findings from the interviews the conceptual framework is developed that shows the cause and effect relationships between all the factors. Also other drivers derived from literature are added to complete the framework. As shown in chapter 2, iterations are performed to optimize the quality of the conceptual framework. The drafts are shown in appendix D. Figure 5 shows the final conceptual framework. It takes the benefits and challenges proposed in section 4.1.3 into account. The preconditions that are identified have to be met at all times in order to gain value from the investment.
Figure 5: Conceptual framework for applying Big Data in the financial services industry
The left side of the conceptual framework shows the solutions that have to be implemented to ultimately gain value. The other boxes and arrows show how to get from the solutions to the value. The lines are independent of each other which mean that reaching the super goal is already possible with meeting one of the sub goals. The conceptual framework contains only factors that are directly related to Big Data in the financial services industry. It is possible that the factors that are provided in this framework conflict with other goals of an organization, but this is excluded from this research.

The solutions that should be implemented to gain value from an investment in Big Data in the financial services industry are installing the hardware and software that Big Data requires, performing analytics and visualization techniques that are possible with Big Data, creating a project team to create culture, methods, and provide training, and attracting new employees that fit best with a Big Data environment. It is important that the new hardware and software are used in combination with the existing IT landscape to be efficient and cost-effective. Moreover, it is important that the new processes derived from using Big Data techniques are synchronized with current processes to make sure that processes could be executed without errors. It is also important that the right analytics and visualization techniques are used to get the right information from the enormous amounts of data that is visualized in such a way that it is usable for the management. Additionally, it is also important that employees and management change the mindset to a new way of working with Big Data and that they are all skilled to use the new technology. To accomplish this it might be necessary to attract new employees and dismiss employees that are not capable to change to the new way of working. When employees are trained to work with Big Data and master it, they will probably be more satisfied. This will lead to a decrease in employee costs because employees are less likely to be absent when they are satisfied with their job. The relation between attracting new employees, employee satisfaction, and decrease in employee costs is the only relation that is not shown in literature but only based on information from the interviews. All the other relations are supported by the document study as well as the interviews with experts.

Implementing these solutions together will provide better information about the main processes in the financial services industry which are payments, savings, financings (credits, lending, and mortgages), investments, and insurances. More complete information can be used to make decisions about these processes. Moreover, when it is necessary, real time or near real time information can be accessed by the use of Big Data which decreases the waiting time on information and improves the accuracy of decisions that have to be based on real time information. These more complete insights will lead to, on the one hand, an increase in better targeted marketing (cross selling, customized deals) and better insights in the competing market, and on the other hand it also decreases wrong credit scoring and stimulates fraud detection. When customers are better targeted with marketing it is likely that the sales, the retention of customer, and the satisfaction of customer will increase because the customer gets an individual treatment. Moreover, better insights in the competing market can lead to an increase in the
market share because an organization now has information on how to improve the organization to outperform the competition. All together this will lead to an increase in revenue. At the cost side, the decrease in wrong credit scoring will decrease the overall costs because credits are provided based on real insights in behavior and not on predetermined categories. Together with an increase in fraud detection and a decrease in employee costs the overall costs will decrease. The increase in revenue and decrease in costs will eventually lead to a positive cash flow generation.

The preconditions that have to be met are privacy, compliance to rules and regulations, and security. First of all, personal information should only be used when rules are not trespassed and the person accepts that the organization uses the information. Second, an organization should be compliant to the rules and regulations that are defined. Especially in the financial services industry there are many different rules which change constantly. Last, the security of information should be optimal at all times to make sure it will not be used by non-authorized persons.

The three V’s (volume, velocity, variety) that are used to define Big Data in general are indirectly shown in the conceptual framework. The enormous amounts of data that characterize Big Data can only be dealt with by Big Data techniques such as Hadoop, MapReduce or other packages. That is why this technology needs to be installed and synchronized with the current technology. Variety is encapsulated in the analytics part. With the use of Big Data it possible to analyze structured and unstructured data to provide information about the financial behavior of customers. The velocity of Big Data is mainly of influence on taking opportunities as fast as possible. Using Big Data technology makes it possible to get (real time) information faster which is mainly important to detect fraud and to provide instant credit scoring.

Business Intelligence has a relation to Big Data as shown in chapter 3. Therefore it is highly likely that the same framework would be constructed when the value of an investment in Business Intelligence in the financial services industry would be researched. However, the difference will be in the intensity of the factors that are provided because of the difference in volume, variety, and velocity of the data. For example, the factor “providing (real-time) information about payments” would still be the same, but the information gathered using Big Data will be more complete than the information gathered using Business Intelligence. Moreover, implementing Business Intelligence packages will be easier because the already existing relational databases will be used.

4.2.1 Small illustration of the use of the framework
To demonstrate the applicability of the framework, this paragraph provides an example about a credit approval process. The credit approval process starts with an invoice of a client to apply for a loan. The financial institution (often a bank) will decide if the application will be approved or not. It depends on the amount that is applied for if one or multiple approvers are necessary to approve the application. At every approval level, the decision is based on the following factors. First of all the credit worthiness will be evaluated. The credit history and expected performance
will be checked to define a debtor’s ability to pay. Second, the size of the debt will be investigated to see if the earnings exceed the demand of the payment schedule. The total debt service ratio determines the debtor’s capacity to handle the debt. Third, the loan size is taken into account to see if the creditor is able to provide the sum of money that is applied for. Fourth, the frequency of borrowing is checked to identify the reputation a debtor already has at the specific financial institution. Moreover, the length of commitment is taken into account because the longer the commitment, the higher the risks. Last, social and community initiatives are considered which make it possible that the creditor accepts the application for the social good (Biz Loan Link, 2013; Helbekkmo, 2006; Inc., 2013; RBCRoyalBank, 2013). Figure 6 shows the credit approval process.

![Credit approval process diagram](image)

*Figure 6: Credit approval process*
The credit approval process is used as input for the conceptual framework to discover if Big Data could be useful in (certain parts of) this process. When Big Data is installed correctly and employees are skilled to work with the technology it is possible to gain more complete insights in the credit history and the current financial situation of the debtor (left side of conceptual framework). Additionally, based on the behavioral patterns of the debtor it is possible to predict if the debtor is likely to repay the complete loan. This will increase the accuracy of making the right decision (box decrease wrong credit scoring) and increases the possibility to provide a customized deal. Moreover, with Big Data it is possible to gain these insights faster which make the credit approval process shorter. This is important for the debtor because the debtor most likely wants the credit as fast as possible. Shortening the process and providing a customized deal will increase the satisfaction of the debtor and the debtor will probably come back again if another loan is necessary. The accuracy of the decision together with the satisfaction and retention of the debtor will decrease the overall costs and increase the revenue that represent the creation of value. The path that is followed in the conceptual framework is shown in figure 7.

---

**Figure 7: Value from Big Data for the credit approval process**
5. Case Study Validation
As explained in the methodology, Hevner et al. (2004) state that evaluation or validation is a crucial component of doing research. For this research work the observational validation method case study is chosen. In this case study, various experts are questioned about the same problem via a questionnaire. The experts are different from the experts that are interviewed. The questionnaire should provide the empirical evidence that validates the conceptual framework. First the method that is followed is described including the process, the respondents, and the reliability and validity. Additionally the results of the questionnaire and the revised framework are shown. Last, a comparison is made with other frameworks to provide context.

5.1 Method of questionnaire

5.1.1 Execution of questionnaire
The questions of the questionnaire are based on the conceptual framework. Every question represents a line of the conceptual framework, indicating to what extent the respondent agrees with the linkage. All questions are presented with the key design issue of capturing the value of Big Data in the financial services industry. The questions are closed questions with the answering options based on a five point Likert scale going from ‘strongly disagree’ to ‘strongly agree’. The five point Likert scale is an interval scale at which the differences in the responses between any two points on the scale remain the same (Sekaran, 2003). Additionally, personal information was requested to have insights in the type of respondents. At the end of the questionnaire it was possible to provide comments.

The questionnaire is published on the internet and accessible via a URL. Attention to the questionnaire was drawn by promoting it on the internet via social media such as LinkedIn, Twitter, and Yammer. The message was only posted in the groups that have affinity with Business Intelligence and/or Big Data and the financial services industry. Accordingly, a mass mail was sent to employees of Capgemini that also have affinity with Business Intelligence and/or Big Data in the financial services industry. Also personal mails were sent to persuade persons to complete the questionnaire. The questionnaire was available for 1,5 months.

Eventually 51 persons responded from which 19 completely filled in the questionnaire and eight partially completed. To analyze the data, the dataset has to be complete. To check if the missing data is random or systematic, the Little’s MCAR $\chi^2$ test is performed from which the results are shown in appendix E1. The significance of the test is 1,000 which is bigger than 0,05 which means that the missing data is completely at random. Therefore it is only possible to use regression based methods to fill in the missing values (Field, 2009). However, first the cases with more than 15% of missing values have to be deleted because completing them would cause too much bias (Field, 2009). Since the questionnaire contains 69 variables, cases with more than 10 missing values have to be deleted. This resulted in a deletion of 24 cases, leaving 27 cases of which 8 are not complete. These 8 cases are completed using the regression based method mean substitution. This method replaces the blank spots by the mean of that variable based on the
answers of the other respondents. Ultimately, a dataset of 27 complete cases is constructed that is used for the analysis. The descriptive statistics of the dataset can be found in appendix E2.

5.1.2 Characterization of respondents
Respondents are asked to provide information about certain demographics. The demographics that are requested are gender, age, nationality, education, current sector, current company, job status, and number of years worked in current organization. The frequencies and percentages are shown in appendix E3.

Gender
The options for gender are male and female. From the 27 respondents there are 21 males (77,8%) and 6 females (22,2%). This indicates that the males have the majority.

Age
The options for age are 0-20, 20-35, 36-50, 51-65, and older than 65. From the 27 respondents, 12 belong to the category 21-35 (44,4%), 11 belong to the category 36-50 (40,7%), and 4 belong to the category 51-65 (14,8%). This division relates to the range of the working population.

Nationality
The option for nationality was open to the respondent to fill in the name of a country. From the 27 respondents, 20 are Dutch (74,1%), 6 are Indian (22,2%), and 1 is Norwegian (3,7%). So, the dataset mainly exist of Dutch people. It does not provide a global representation.

Education
For this demographic it was asked what the highest completed level of education is. The options ranged from high school to a master degree. It was also possible to fill in another answer. For the 27 respondents, 3 completed high school (11,1%), 9 achieved a bachelor degree (33,3%), and 15 accomplished a master degree (55,6%). This means that 88,9% belongs to the category higher education, which is not a representation of the working population.

Current sector
The answering options for the sector the respondent is currently working in were financial services, retail, care, and the option to fill in another sector. For the 27 respondents, 20 currently work in the financial services industry (74,1%), 2 in retail (7,4%), 1 in care (3,7%). The 4 other respondents all filled in IT consulting as the sector in which they currently work (14,8%). This indicates that the majority has knowledge of the financial services industry at which this research work is focused.
Current company

In the questionnaire it was asked to fill in the name of the current company the respondent is working. For the 27 respondents, 21 work currently at Capgemini (77,8%) and 6 work somewhere else (22,2%). The other companies are Volkswagen Pon, Rivierduinen, NN, Rabobank, ASR, and ABN Amro.

Job status

Four options were provided to indicate the job status of the respondent: first level supervisor, middle management, top management, and nonmanagerial. For the 27 respondents, 2 are first level supervisor (7,4%), 5 are in middle management (18,5%), and 20 have nonmanagerial jobs (74,1%). This means that top management is not represented in this dataset.

Number of years worked in the current organization

The last demographic that was asked for is the number of years that the respondent works at the current organization. The options were less than 1, 1-2, 3-5, 6-10, and more than 10 years. For the 27 respondents, 1 works at the current organization less than 1 year (3,8%), 10 work 1-2 years at the current organization (37%), 4 work 3-5 years at the current organization (14,8%), 4 work 6-10 years at the current organization (14,8%), and 8 work more than 10 years at the current organization (29,6%).

5.1.3 Reliability and validity

Besides dealing with missing data, also the reliability and validity of the data has to be checked to test the goodness of the data.

The reliability indicates the degree to which the measurements give the same result. Testing the reliability provides information about the consistency of the data (Sekaran, 2003). A way to test the reliability of the data is with the reliability coefficient Cronbach’s Alpha. This coefficient shows how well items in a data set are positively related to one another. The value of the Cronbach’s Alpha can be between 0 and 1. The higher the value, the more reliable the data is (Field, 2009; Sekaran, 2003). The formula for the Cronbach’s Alpha (α) is:

$$\alpha = \frac{k \cdot (\frac{\text{covariance}}{\text{variance}})}{1 + (k-1) \cdot (\frac{\text{covariance}}{\text{variance}})}$$, $k = \text{number of items}$

As shown in table 3, $\alpha = 0.983$ which means that the data is very reliable. Moreover it was tested if the reliability would go up when a variable would be deleted, but this was not the case (Appendix E4).
Table 3: Cronbach’s Alpha

<table>
<thead>
<tr>
<th>Reliability Statistics</th>
<th>Cronbach's Alpha</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.983</td>
<td>69</td>
</tr>
</tbody>
</table>

The validity of data represents if a measure assesses the construct that it is intended to measure (Field, 2009; Sekaran, 2003). There are different kinds of validity, but the two must used ones are convergent validity and discriminant validity. Convergent validity requires a high degree of correlation between two different sources responding to the same measure. Discriminant validity requests that two distinctly different concepts should not be correlated to each other (Field, 2009). To test both types of validity a correlation matrix is used that provides the strengths of a relationship between two variables without providing a direction of the relationship. Because the data set contains only 27 cases, it is not possible to assume that the data is normally distributed. Therefore normality tests are done to test the normality of the data. Appendix E5 shows examples of the normality tests including P-P plots and histograms for every variable. A P-P plot is a probability-probability plot that plots the cumulative probability of a variable against the cumulative probability of the normal distribution. When the values are close to the diagonal, it can be assumed that data is normally distributed (Field, 2009). A histogram shows if the line of distribution is the same as in the plot of the normal distribution. From the results of these tests it can be concluded that it cannot be assumed that the data is normally distributed. Therefore it is chosen to use Spearman’s correlation coefficient to check the validity of the data, because this coefficient does not require data to be normally distributed. Because the proposed relationships between variables are directional, a one-tailed test is used. The coefficients are in the range of -1 to 1, in which -1 means a perfect negative relationship, 0 means no relationship, and 1 means a perfect positive relationship (Field, 2009). Some examples of the correlation matrix are shown in appendix E6.

Looking at the correlation matrix, it seems that variables that should have a strong relationship correlate highly and variables that should not have a relationship correlate on a low level. Though, three correlations are striking. First, the processes payments, savings, financings, investments, and insurances correlate highly with each other for variables such as customized deals and synchronizing IT, often a correlation of 0.9 or higher. This is not a problem for the validity, however it could be a signal that the processes should be combined into one variable. The second correlation that is noticeable is the relation between cross-selling and retention of the customer, which is quite low while it was expected to be high. The last correlation that is lower than expected is the relation between the employee costs and costs in general. Although these last correlations indicate that the discriminant validity is affected, overall the correlations do not show many deviations. Therefore, the requirements of convergent and discriminant validity are met.
5.2 Results of questionnaire

As said in the previous paragraph, it cannot be assumed that the data is normally distributed. Therefore it cannot be assumed that the data set is a representation of the population, which makes it difficult to perform any regression analysis (Field, 2009). For this research work, descriptive techniques are used to give an indication about to what extent the respondents agree with certain relationships. The correlation matrix is a descriptive technique, however it does not provide any direction of the relationship which makes it difficult to use only this technique. Because all answers are provided using a five point Likert scale, it is possible to directly compare the answers. To provide an insight in the strength of the relationships based on the opinion of the respondents, weights are calculated for every line. For every variable the mean is divided by five to get a relative weight between 0 and 1 that can be compared directly. A weight of 0-0,3 means that the respondents are strongly disagreeing with the initial relationship. A weight of 0,3-0,5 means that the respondents disagree with the relationship. A weight of 0,5-0,7 means that the respondents are neutral to the relationship. A weight of 0,7-0,9 means that the respondents overall agree with the relationship. Last, a weight of 0,9-1 means that the respondents strongly agree with the relationship. For difficult cases it is possible to use the correlation matrix to see if a relationship should be removed or not.

Calculating these weights show that on the overall the respondents slightly agree with the relationships claimed in the conceptual framework. The relationships that are noticeable are shown in table 4 and 5. Table 4 shows the relationships that are perceived as strong (weight is greater than 0,75). Table 5 shows the relationships that are perceived as neutral or less. The relationships with a weight lower than 0,65 are in bold in table 5. The complete overview of the weights of the relationships is shown in appendix E7.

Table 4: Strong relationships (weight > 0,75)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase revenue</td>
<td>Increase EBIT</td>
<td>0,793</td>
</tr>
<tr>
<td>Increase fraud detection</td>
<td>Decrease costs</td>
<td>0,807</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase customer satisfaction</td>
<td>0,762</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase retention of customer</td>
<td>0,763</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase sales</td>
<td>0,756</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Increase fraud detection</td>
<td>0,785</td>
</tr>
<tr>
<td>Visualization and analytics of data</td>
<td>Information about payments</td>
<td>0,778</td>
</tr>
<tr>
<td>Visualization and analytics of data</td>
<td>Information about savings</td>
<td>0,770</td>
</tr>
<tr>
<td>Visualization and analytics of data</td>
<td>Information about financings</td>
<td>0,785</td>
</tr>
<tr>
<td>Visualization and analytics of data</td>
<td>Information about investments</td>
<td>0,793</td>
</tr>
<tr>
<td>Visualization and analytics of data</td>
<td>Information about insurances</td>
<td>0,785</td>
</tr>
</tbody>
</table>
Table 5: Relationships that are considered neutral or less (weight < 0.7)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase market share</td>
<td>Increase revenue</td>
<td>0.674</td>
</tr>
<tr>
<td>Increase cross selling</td>
<td>Increase retention of customer</td>
<td>0.685</td>
</tr>
<tr>
<td><strong>Decrease employee costs</strong></td>
<td><strong>Decrease costs</strong></td>
<td><strong>0.585</strong></td>
</tr>
<tr>
<td>Information about financings</td>
<td>Improve insights in competing market</td>
<td>0.681</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Improve insights in competing market</td>
<td>0.689</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Decrease wrong credit scoring</td>
<td>0.674</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase fraud detection</td>
<td>0.689</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase fraud detection</td>
<td>0.667</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase cross selling</td>
<td>0.689</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase cross selling</td>
<td>0.696</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase customized deals</td>
<td>0.674</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase customized deals</td>
<td>0.696</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about payments</td>
<td>0.644</td>
</tr>
<tr>
<td><strong>Provide training in use new IT</strong></td>
<td>Information about savings</td>
<td><strong>0.644</strong></td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about financings</td>
<td>0.652</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about investments</td>
<td>0.654</td>
</tr>
<tr>
<td><strong>Provide training in use new IT</strong></td>
<td>Information about insurances</td>
<td><strong>0.644</strong></td>
</tr>
<tr>
<td>Provide training in business changes</td>
<td>Decrease employee costs</td>
<td>0.656</td>
</tr>
<tr>
<td><strong>Increase employee satisfaction</strong></td>
<td><strong>Decrease employee costs</strong></td>
<td><strong>0.600</strong></td>
</tr>
<tr>
<td>Attract new employees</td>
<td>Increase employee satisfaction</td>
<td>0.616</td>
</tr>
<tr>
<td>Create project team</td>
<td>Increase employee satisfaction</td>
<td>0.640</td>
</tr>
<tr>
<td>Create project team</td>
<td>Provide training in business changes</td>
<td>0.669</td>
</tr>
<tr>
<td>Attract new employees</td>
<td>Provide training in use new IT</td>
<td>0.685</td>
</tr>
<tr>
<td>Installing Big Data packages</td>
<td>Synchronize IT</td>
<td>0.692</td>
</tr>
<tr>
<td>Installing Big Data packages</td>
<td>Combine legacy and new hardware and software</td>
<td>0.669</td>
</tr>
</tbody>
</table>

To discover if certain relationships of the conceptual framework should be removed, links with a weight lower than 0.65 are investigated using the correlation matrix (shown in bold in table 4). Based on the correlations together with the weights there is no evidence that there will be a decrease in employee costs that decreases the overall costs. Moreover, there is no evidence that employee satisfaction has an influence. However, there is evidence that training will lead to a decrease in overall costs. Additionally, no comments were provided at the end of the questionnaire that have an influence on the framework. Together this leads to the final framework, *Big Data Value Framework for the Financial Services Industry*. The parts that are changed compared to the conceptual framework are shown in figure 8 and the complete framework is shown in figure 9. It is chosen to show the different processes in the financial services separately although some relationships could be combined based on the correlation matrix. This is done to be clear that there are different processes and that the strength of the relationships are not the same. The list of weights of the final framework is shown in appendix F.
Figure 8: Differences between conceptual framework and Big Data Value Framework for the Financial Services Industry

Figure 9: Big Data Value Framework for the Financial Services Industry
5.3 Comparison of the different frameworks

As discussed in chapter 3, literature lacks information about a value framework for Big Data in general and for the financial services industry as well. With the development of the *Big Data Value Framework for the Financial Services Industry*, a start is made to cover this gap. A comparison with other value frameworks for Big Data could provide perspective on the disposition of this framework in literature, however, there are no other frameworks yet identified. Therefore it is only possible to compare the general frameworks to gain insight in the disposition. There are many frameworks related to performance management or value management. One of the frameworks is the Performance Pyramid of Lynch and Cross (1995). This framework links performance measures at different levels in an organization by combining an organization’s strategy with operations. The framework is focused on internal efficiency and external effectiveness. Comparing this framework with the Benefit Logic Model, it is clear that both frameworks incorporate business strategy and business operations. Moreover, both frameworks provide overall indicators, but do not provide any support how to measure these indicators. A difference between the frameworks is that the Benefit Logic Model identifies the indicators, while the Performance Pyramid framework does not. Another difference is that the Benefit Logic Model is focused on revenue and costs, while the other framework focuses on efficiency and effectiveness.

Another performance framework which is often used in practice, is the Balanced Scorecard developed by Kaplan and Norton (1996; 2000). This framework is more elaborate than the Performance Pyramid because it provides information to identify measures and it pays attention to the issue of continuous improvement. The Balanced Scorecard complements financial measures of past performance with the measures of the drivers of future performance in which the objectives and measures are derived from strategy. The Balanced Scorecard is divided into four perspectives that all incorporate a perspective from which organizational performance can be viewed: financial, customer, internal business process, and learning and growth. All perspectives are divided in objectives, measures, targets, and initiatives. Comparing this framework with the Benefit Logic Model results in some similarities and differences. A similarity between the frameworks is that they both incorporate different perspectives from which value can be derived. Moreover they both identify objectives that have to be realized. A difference between the frameworks is that the Balanced Scorecard framework takes a step further and provides possible measures to measure the performance of the objectives, while the Benefit Logic Model only shows the objectives itself. Figure 10 shows the *Big Data Value Framework for the Financial Services Industry* divided into the four perspectives of the Balanced Scorecard to provide an example of how the *Big Data Value Framework for the Financial Services Industry* could be compared to a Balanced Scorecard that is focused on Big Data.
Figure 10: Big Data Value Framework for the Financial Services Industry divided into four perspectives of Balanced Scorecard
As the figure shows, not all factors could be accommodated to the four perspectives. The two factors that do not fit in the four perspectives contain the technical aspect of the framework. Kaplan and Norton (1996; 2000) place the technical aspects into the learning and growth perspective which is divided into two subcategories: employee capabilities and empowerment, and the information system capabilities. Although Kaplan and Norton (1996; 2000) combine them together into the learning and growth perspective, it might be better to separate this perspective into two perspectives, because they are distinct from each other and have a different influence on the other perspectives. This will clarify to organizations which perspective should be improved and which perspective is already correct.

A last framework that is interesting to compare the Benefit Logic Model with is the Performance Prism framework from Neely et al. (2002). The Performance Prism tries to integrate five related perspectives to provide a structured business performance model. The five related perspectives are stakeholder satisfaction, stakeholder contribution, strategies, processes, and capabilities. The Performance Prism is based on the Balanced Scorecard, but takes all stakeholders into account while the Balanced Scorecard only focuses on the shareholder which is often criticized in literature. The Performance Prism is proposed to show the complexity of organizations’ relationships with the multiple stakeholders within the context of the operating environment, e.g. customers, employees, suppliers, government, and investors. This focus on all stakeholders is also not incorporated in the Benefit Logic Model, which is mainly focused on gaining value for the shareholders. A similarity between the frameworks is that they both show different perspectives from which value can be derived.

In this research work only a value framework for the financial services industry is developed. It is possible to use the Big Data Value Framework for the Financial Services Industry as a reference to develop a Big Data value framework for other industries such as logistics, manufacturing, retail, or care. The main difference between these industries is the various processes that are specific for an industry. The factors related to the processes have to be adapted to create a correct framework for a different industry. The technical implementation, the cultural change, the influence on customers (other businesses or persons), and the financial consequences are in essence the same for every industry. Only the strength of a certain factor and the relationships will be different across industries. Therefore it is possible to use the developed framework as a starting point to construct a Big Data value framework for another industry. To note, attention have to be paid to the description of the different factors and the relation between them. The difficulty of using the Benefit Logic Model is that objectives are only described and not quantified in numbers which makes it receptive for different interpretations. Moreover it is difficult to determine if all factors and their relations are included and correct to assess the value creation of Big Data. On the other hand, the Benefit Logic Model provides a concrete overview of the cause and effect relationships between different factors and identifies how to create value from a certain investment. That is why the Benefit Logic Model is especially useful for research that is not yet matured.
The value framework of Big Data shown in this research work is a framework that is not part of an overall business strategy or business model. However, it could be used to create new business models or adapt existing business models. For example it could be used for the building block value propositions from the Business Model Canvas proposed by Osterwalder and Pigneur (2009). Moreover, it could also be used as a first step in the value activity viewpoint from the e3 Value Model of Gordijn (2004). In this way the value framework will be extended and some context will be provided.
6 Conclusion, Limitations, and Future Work

6.1 Conclusion

This section provides answers to the research questions described in chapter two. The first research question, **What is Big Data?**, refers to the definitions of Big Data together with the positive and negative aspects. As emphasized in section 3.1.2, a general accepted description of Big Data is based on the three V’s volume, velocity, and variety. These V’s separate Big Data from other trends. Different sources claim that there are also a fourth and fifth V (veracity and value), but these are not generally accepted. However, this research shows that Big Data can provide a certain value. With Big Data it is possible to retrieve large amounts of data relatively quickly which was not feasible before.

Using Big Data provide several benefits on an operational, financial, and business level. One of the ultimate benefits of using Big Data is that an organization can have quicker access to cleaner, more relevant data to drive insights and optimize decision making. Other benefits that relate to this ultimate benefit are a more complete understanding of market conditions and evolving business trends, more accurate responses to customer needs, consistency of decision making, better risk management, and a competitive advantage. Big Data can also help organizations to comply with the increasing complex and always changing regulatory requirements.

Having the right people with the right skills and a right organizational structure is one of the difficulties of implementing Big Data. Many organizations have more trouble with the managerial and cultural barriers rather than data and technology barriers. One of the major complications for successfully implementing Big Data is talent. Moreover, the culture of the company should be changed to a culture in which the use of Big Data is more valued and rewarded. There are also technological difficulties. Dealing with the volume of data is often not a problem for organizations, the difficulty lies in the velocity and variety of data. Today’s relational database architectures need to be modernized to adapt to the increasing volume, variety, and velocity without removing the existing systems which is a complicated process. Additionally, another difficulty is to integrate all information from different data silos together without separating it again in different silos. A more general difficulty that plays a role when using Big Data is the issue of privacy. Organizations can only use personal information if it is allowed by the rules and if a customer accepts that the information will be used.

The second research question concerns the relationship between Business Intelligence and Big Data. As shown in section 3.1, Business Intelligence and Big Data both represent the overall technology and business package that transforms data into information and knowledge to ultimately create value for the organization. Business Intelligence is composed of a technical and a managerial viewpoint. From a technical point of view, Business Intelligence is the set of tools that supports the process described by the managerial viewpoint. It follows the process of gathering data from different systems, placing the data in a data warehouse, transforming the data, analyzing the data, and providing information to improve decision making. The managerial
viewpoint sees Business Intelligence as a process in which data from inside and outside the organization is integrated to generate relevant information for the decision-making process. This managerial view is more related to the term Business Information Management. Business Information Management encompasses the management process of converging data into information to gain insights and foresights in business performance. Both gaining insights and foresights are boosted by the use of Big Data. Just as Business Intelligence Big Data still focuses on providing information extracted from data to enhance the decision making processes, however, with Big Data it is also possible to use unstructured raw data for the analysis. Moreover, the data that is used is bigger, has more variety, and goes a lot faster. Whereas Business Intelligence retrieves information from Relational Database Management Systems, Big Data tries to combine data from different management systems together to provide more accurate information. In this way it is possible to retrieve a large amount of data relatively quickly which makes it possible to provide real-time information instead of information from the past.

Addressing the third research question regarding **the most relevant approaches to assess the business value of Big Data in an IT context**, business value can be defined in many ways. As indicated in section 3.2, business value can be described as everything that adds value to the organization. Specializing it to IT, business value can be described as the organizational performance impacts of IT at both the intermediate process level and the organization wide level, and comprising both efficiency impacts and competitive impacts. This also holds when the business value of Big Data is defined in which IT will be replaced by Big Data. The business value of Big Data can be summarized as making the organization more intelligent by making decisions based on good information. The business value of Big Data is captured in the possibility to analyze information that could not be used before and addressing the storage and real time accessing problems at once.

Multiple approaches could be used to construct business value. A comprehensive method that incorporates all possible measures of business value is the use of frameworks. Different performance management frameworks could be used to define the business value of a certain investment such as the quantitative frameworks Performance Pyramid, Balanced Scorecard, and Performance Prism as well as the more qualitative frameworks such as the Business Model Canvas, e3 Value Model, and the Benefit Logic Model.

The last research question is: **How can the business value of Big Data be determined in the area of financial services?** As stated in chapter four and five, the business value of Big Data in the financial services industry can be described as everything that adds value to the performance of a financial organization when Big Data is used. To show how business value can be created, the Benefit Logic Model is used to create a value framework for investing in Big Data in the financial services industry. Solutions that should be implemented are installing the hardware and software that Big Data requires, performing analytics and visualization techniques that are possible with Big Data often done by a data scientist, making a project team to create culture, methods and provide training, and attracting new employees that fit best with a Big Data
Together with drivers such as customized deals, cross selling, and fraud detection this will lead to an increase in revenue and decrease in costs. However, the preconditions privacy, security, and compliance to rules and regulations have to be met at all times to create business value.

Comparing the *Big Data Value Framework for the Financial Services Industry* with other frameworks could provide the disposition of the framework in literature, however no other frameworks are available yet. Therefore it is only possible to compare the general performance frameworks. All frameworks have the similarity that business strategy and business operations are incorporated. A difference between the frameworks is that some frameworks do not provide objectives or indicators (Performance Pyramid), while others provide objectives and measurements to measure these objectives (Balanced Scorecard). Another difference is that frameworks such as the Performance Prism include all stakeholders involved, while the Benefit Logic Model is mainly focused on the shareholders. To take these differences into account it is possible to incorporate the *Big Data Value Framework for the Financial Services Industry* into the creation of new business models or adaption of existing models in which all parts of an organization are addressed.

6.2 Limitations
Although this research is conducted with great care, still some limitations can be identified. First of all, Big Data and especially the value of Big Data in the financial services industry is a subject that is not researched extensively yet. To get an overall impression about the subject also less reputable magazines are used. On the other hand, new articles are published almost every day because Big Data is quite new and a hot topic at this moment. That is why it is difficult to keep track of all the information that is published. Therefore it is possible that the latest information is not incorporated in this research work. Some articles that are used could be outdated, even when they are only published last year.

A second limitation of this research work is the small number of interviews and low response on the questionnaire. Only eight interviews are performed which provide insights in the opinions of people who work at organizations, but the number is too low to generalize the information. This also holds for the information collected via the questionnaire. With only 27 respondents it is not possible to construct a value framework that could be generalized. That is why the *Big Data Value Framework for the Financial Services Industry* is only usable for this particular sample of the population. The framework could be used as an indication which variables have an influence on the business value of Big Data, though it is not completely proven.

Another limitation that is closely related to the previous one is that it could not be proven that the data set is normally distributed, which made it only possible to use descriptive analyzing techniques. When the data set is normally distributed it is possible to use regression techniques that provide information that could be generalized for the complete population. Now it was only possible to describe what the respondents’ opinions are.
A fourth limitation is the lag of the development of drivers at the cost side of the *Big Data Value Framework for the Financial Services Industry*. From the interviews and questionnaire mainly drivers for the revenue side were found, whereas the drivers of the cost side lagged behind. This might be caused by the respondents’ focus on revenue instead of costs. For Big Data it also might be easier to indicate the drivers for revenue than for costs.

The last limitation for this research work is that with the setup of this research it was only possible to test influences of factors directly linked to each other. It was not possible to discover influences of factors that were not directly linked in the conceptual framework, although there might have been other linkages.

### 6.3 Recommendations and directions for further research

This research work is a first step in the development of a value framework of investing in Big Data in the financial services industry. One of the recommendations is to use the *Big Data Value Framework for the Financial Services Industry* as a starting point for developing a business case in the financial services industry. Another recommendation is to use the framework as a support to explain how value can be created to organizations that doubt about investing in Big Data. A last recommendation is to develop the framework into a framework that provides specific measures for each variable to measure if investing in Big Data indeed creates value. This could be measured by first measuring the predefined variables before Big Data is implemented and then measure again when Big Data is implemented and used for a while. When the performance indicators are better, it could be said that Big Data indeed creates business value.

A direction for further research is to expand this research by collecting more data to make it possible to generalize the results of the research. Another direction for further research is to expand this research by investigating if there are more factors and if different factors are linked together that not have been investigated in this research work. The focus of next explorations could be on the cost side of the framework, because that side is underexposed in this research work. Both directions might provide new insights in the creation of business value of Big Data in the financial services industry.

Another direction for further research is to examine if it is possible to define indicators for the factors found during this research. With these indicators it is possible to measure the performance of the organization when it uses Big Data. Also targets have to be identified to compare the outcome of the indicators with standardized targets that have to be met. In this way it is possible to quantify the business value.

A last direction for further research is to further develop this framework by incorporating for example actors that are involved and the interests they have. The framework could be used as a starting point to develop new and improve existing business models that represent a part of the organization.
References


Bean, R. (2013, January 01). Organizational alignment is key to Big Data success. (D. Kiron, Interviewer)


Appendices

Appendix A: Interview questions

General

- What includes your current job function?
- For how long do you work at the current organization?
- To what extent do you work with the concept Big Data?

Business Intelligence

- What is your definition of Business Intelligence?
- What are the core elements of Business Intelligence?

Big Data

- What is your definition of Big Data?
- What are the core elements of Big Data?
- What are the advantages and disadvantages of the use of Big Data?
- To what extent is Big Data already used in practice?
- What are the similarities and differences between Business Intelligence and Big Data?
- To what extent is it possible to use knowledge about Business Intelligence for Big Data?
- To what extent is it possible that Big Data provides new insights in information needs compared to Business Intelligence?

Value Business Intelligence and Big Data

- What is your definition of business value?
- How is it possible to measure business value?
- Is it possible to measure the business value of Business Intelligence?
- Is it possible to measure the business value of Big Data?
- Is it possible to use the same measuring methods to measure the business value of Business Intelligence and Big Data? Are there any differences?
- Which measuring methods do you think are appropriate to measure the business value of Big Data?
- What are the top three drivers for business value of Big Data?
- What are the top three success factors for the business value of Big Data?
- Which variables should definitely be measured concerning the business value of Big Data?
- Is it possible that Big Data creates business value in the financial services industry?
- In which parts of the financial services industry does Big Data provide business value?
- Which variables should definitely be measures concerning the business value of Big Data in the financial services industry?
Appendix B: Contact details
In case of any questions concerning this document, the interviews, or questionnaire please contact via the details provided below.

Anke de Vries

h.a.d.vries@student.tue.nl

Appendix C: Detailed information chapter 3

The relation between Business Intelligence and Big Data

Business Intelligence, Business Information Management and Big Data represent the overall technology and business package that transforms data into information and knowledge to ultimately create value for the organization. The next paragraphs provide information on the different subjects and their relationship.

Business Intelligence

Definition
The first term that comes across in history about using data intelligently is Business Intelligence. In literature there are different, semi-identical definitions used for the term Business Intelligence. The term is used for the first time in a paper of Luhn (1958) who developed a Business Intelligence System to “utilize data-processing machines for auto-abstracting and auto-encoding of documents and for creating interest profiles for each of the "action points" in an organization” (p. 314). Business refers to a collection of activities exercised for whatever purpose. Intelligence refers to “the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal” (Luhn, 1958, p. 314). In the 1960s the term Business Intelligence was related to the decision support systems that were developed. The term evolved over the years and nowadays Business Intelligence is used as a replacement for management information systems, and executive information systems and includes decision support, data mining, online analytical processing (OLAP), querying and reporting, statistical analysis, and forecasting (Thomsen, 2003). Business Intelligence as an umbrella term is introduced in 1989 by H. Dresner of the Gartner Group who described it as ‘all the technologies that help business make decisions based on fact’; ‘Using fact rather than intuition was the key to intelligence’ (Nylund, 1999).

According to Olszak and Ziemba (2003) Business Intelligence systems consist of different modules. The systems contain tools to extract and transfer data, data warehouses to store data, analytic tools (OLAP) to access, analyze and model the data, tools to report the analysis, and a presentation layer that provides information to users in a comfortable way. To adequately design, implement, and use Business Intelligence systems, four dimensions are taken into consideration. The business dimension includes selection of management methods and techniques. The
functional dimension involves the function determination of the systems. The technological dimension is based on the selection of information tools, methods, and solutions to build the systems. The organizational dimension contains the methodology determination of the implementation in an organization. Negash (2004) defines Business Intelligence as “combining data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers” (p. 178). A different definition comes from Lönnqvist & Pirttimäki (2006) who divide the term Business Intelligence between a more managerial viewpoint and a more technical viewpoint. The term Business Intelligence can refer to the concept of “relevant information and knowledge describing the business environment, the organization itself, and its situation in relation to its markets, customers, competitors, and economic issues” (Lönnqvist & Pirttimäki, 2006, p. 32). On the other hand, it can also refer to the concept of “an organized and systematic process by which organizations acquire, analyze, and disseminate information from both internal and external information sources significant for their business activities and for decision making” (Lönnqvist & Pirttimäki, 2006, p. 32). Jourdan et al. (2008) support this division by presenting Business Intelligence as “both a process and a product. The process is composed of methods that organizations use to develop useful information, or intelligence, that can help organizations survive and thrive in the global economy. The product is information that will allow organizations to predict the behavior of their competitors, suppliers, customers, technologies, acquisitions, markets, products and services, and the general business environment with a degree of certainty”(p. 121). So, there is a division between a more managerial viewpoint on Business Intelligence and a more technical viewpoint. The managerial viewpoint sees Business Intelligence as a process in which data from inside and outside the organization is integrated to generate relevant information for the decision-making process. From a technical point of view, Business Intelligence is the set of tools that supports the process described by the managerial viewpoint. This division is also implicitly shown in the definition given by Rud (2009) who describes Business Intelligence as a set of methodologies, architectures, processes, and technologies that transform raw data into meaningful and useful information. However, the managerial viewpoint overlaps with the definition of Business Information Management. This is possible because the terms Business Intelligence and Business Information Management are often used interchangeably. In this research work the managerial point of view as described above belongs to Business Information Management and the technical viewpoint describes Business Intelligence.

Different definitions are given, but the main point that is followed during this research work is that Business Intelligence contains the mixture of raw data and analytical tools to present useful information that supports organizations in making decisions. This raw data can be structured or unstructured. Structured data is data that fits neatly into relational or flat files, while unstructured, also called semi-structured, data does not fit decently into relational or flat files. Business Intelligence uses mostly structured data for the analysis of data and delivers the
information to the decision maker (Negash, 2004). Other terms are also used for Business Intelligence, such as Competitive Intelligence, Corporate Intelligence, Competitive Information, or Commercial Intelligence. These terms emphasize the competitive advantage that can be taken by the use of intelligence. According to McGonagle & Vella (2002), Competitive Intelligence consists of two parts: “The use of public sources to develop data (raw facts) on competition, competitors, and the market environment” and “the transformation, by analysis, of that data into information (usable results) able to support business decisions” (p. 36). Five phases can be determined that are all linked together via feedback loops that together describe the Competitive Intelligence cycle. These phases are establishing CI needs, collecting the raw data, evaluating and analyzing the raw data, communicating the finished intelligence, and taking action. These phases could be followed for different reasons. Competitive Intelligence can be strategy-oriented, technology-oriented, tactics-oriented, or target-oriented (McGonagle & Vella, 2002). Negash (2004) described Competitive Intelligence as “a systematic and ethical program for gathering, analyzing and managing external information that can affect your company’s plans, decisions and operations” (p. 186). Although the terms Competitive Intelligence and Business Intelligence are used interchangeably, there is a difference. McGonagle & Vella (2002) describe the difference by discussing that Business Intelligence is only oriented to the internal processes, while Competitive Intelligence is also oriented to the external processes. According to Negash (2004), Competitive Intelligence is a specialized branch of Business Intelligence.

Benefits and challenges
The goal of Business Intelligence is to handle large amounts of information and to identify and develop new opportunities. This can provide a competitive advantage in the market and long-term stability (Rud, 2009). So, Business Intelligence integrates analysis of data with decision-making tools to provide the right information, on the right time, in the right location, to the right persons throughout the organization (Ghazanfari, Jafari, & Rouhani, 2011). In short, it supports organizations to gain business advantage from data. Benefits of using Business Intelligence can be financial related such as cost and time savings. Other benefits are identifying threats and opportunities, understanding the vulnerability of the organization, decrease reaction time, and out-think competition (Thomas Jr., 2001). Willen (2002) shows that Business Intelligence supports in strategic and operational decision making by using it for corporate performance management, optimizing customer relations, monitoring business activity, traditional decision support, and management reporting of business intelligence. In the field of financial services, Business Intelligence systems can be used to improve efficiency and visibility in the back-office operations, because all levels of management can see how problems are related (De Voe & Neal, 2005). Dave (2009) divides the advantages of using Business Intelligence into three categories: revenues, customer relationship, and sales and marketing. Business Intelligence distinguishes the products and services that drive revenues and it ranks customers based on their profitability. Additionally it categorizes low-value customers and tries to improve their value. Moreover, it will detect customer relationship problems early and takes action to resolve them. Lastly, high-value customers are aimed at to reduce the market
risk, successful promotions are ranked based on product and market segment, and it discovers what is in the sales pipeline. Also intangible benefits occur such as the hope that it will lead to a big bang return at some time in the future. However, this does not happen frequently and it is not possible to forecast it (Negash, 2004). Additionally, the databases that are used have to be current, timely, and accurate to benefit from it (Thomas Jr., 2001). Moreover, another disadvantage is that there will be costs involved to operate complete Business Intelligence, such as hardware costs, software costs, implementation costs, and personnel costs. A comprehensive budget should be available to cover these costs from Business Intelligence (Negash, 2004).

**Technology**

Over the years data warehouses became the most important technology for Business Intelligence. The transformation of data into information happens by integrating data from different sources done in a data warehouse (Inmon, 2013). A data warehouse is an independent business database system that is filled with data that is extracted from different sources in order to improve the decision-making process within an organization (Curtis & Cobham, 2002). The data warehouse consolidates and standardizes the data to make it applicable for management analysis and decision-making (Laudon & Laudon, 2012). A data warehouse consists of several components. The Extract and Transform component extracts current and historical data from different internal and external operational systems in the organization. The Information Directory component provides information about the data that are currently in the Data Warehouse. Last, the Data Access and Analysis component gives data access and options for analysis. It is often referred to as the local warehouse in which aggregated data from the Data Warehouse is stored (Laudon & Laudon, 2012).

The process that is followed with Business Intelligence contains four steps. First, data is gathered from different systems and placed into a data warehouse. Second, the gathered data is transformed to make sure that the data is uniform and can be compared. Next, the collected data is analyzed to transform it into information that is usable for the management. Last, the found information is presented via a dashboard, for example to improve the decision making process (Laudon & Laudon, 2012). Kemper et al. (2013) show this process by identifying three layers in the Business Intelligence architecture. The bottom layer, the data support layer, stores structured and unstructured data for decision support. Structured data is often stored in operational data stores, data warehouses, and data marts (smaller data collections of a core data warehouse), while unstructured data is stored using document and content management systems. The middle layer, the information generation, storage, and distribution layer, is used to analyze the data and to distribute relevant knowledge. To analyze the data, Online Analytical Processing (OLAP) and Datamining are used. For the distribution of knowledge, knowledge management systems are provided. The upper layer, the information access layer, provides access to all relevant functions in an integrated environment that is defined by user roles and rights (Kemper et al., 2013).
Big Data

According to Arnold (2012) and Capgemini (2013b) organizations use Business Intelligence applications that provide information to the core of their businesses. However, these applications only represent a small part of the data that can be analyzed. BI tools are good in analyzing structured data while organizations also store many unstructured data that can provide information.

Definition

Quite recently a new term is introduced in Business Information Management that allows systems to transform unstructured data into information: Big Data. Literature and practice do not have a distinct vision about the content of Big Data. On the one hand, Big Data is seen as a new philosophy that is a successor of Business Intelligence, while on the other hand it is defined as a new technology that is part of Business Intelligence (De Vries, 2013b). One generally accepted description of Big Data is based on the three V’s: volume, velocity, and variety (Capgemini, 2012; Database Trends and Applications, 2012; De Vries, 2013b; IBM, 2013; Lopez, 2012; Won, 2013). The volume refers to the enormous amounts of data that are gathered on a daily basis, for example from Twitter. Velocity reflects the speed level at which data is streamed. Variety stands for the numerous sources of data that is streamed into, out of, and through organizations like transactions, social media and websites. Next to that, this streaming happens in a big range of different formats from traditional data to message documents, videos, and audio files.

These three V’s separate Big Data from other trends (De Vries, 2013b). The data volumes are higher than what is processed before, the volumes are larger than what is handled by traditional database technology, external data is added to the existing internal data, various data is combined to analyze, and possibilities are created to gather information near real-time (Capgemini, 2013e). Different sources claim that there is a fourth V: veracity (IBM, 2013). Veracity represents the level of trust that the information on which decisions are made is correct. Some sources also indicate that there might be a fifth V which represents the value that could be created by using Big Data to enable enhanced decision making, to get insight discovery, and to optimize processes (De Vries, 2013b). However, the fourth and fifth V are not generally accepted.

A different description of Big Data is shown in Capgemini (2013e) that defines Big Data by three elements which are the data itself, the process for dealing with data, and the holistic view it can enable. Although there are different definitions, Big Data can be described as “massive amounts of stored content (structured or unstructured) that can be easily analyzed in real time (a reasonable amount of time to get a useful answer)” (Arnold, 2012, p. 32). According to Boyd and Crawford (2012), Capgemini (2013e), and De Vries (2013b), the term Big Data only refers to the quantity of data, forgetting the growing velocity and variety that separates it from Business Intelligence. That is why Boyd and Crawford (2012) define Big Data as an interplay of the technology which maximizes computational power to gather and compare large data sets, the
analysis which identifies patterns to make economic, social, technical or legal claims, and the belief that large data sets offer intelligence with the aura of truth, objectivity and accuracy.

Benefits and challenges

Big Data advances the core technologies needed to collect, store, preserve, analyze, manage, and share enormous amounts of data (Lazar, 2012). Data can be collected across business units, from partners, and customers while a flexible infrastructure integrate all this information effectively. Analytics finally create sense in all of this information (Brown et al., 2011). These predictive analyses are important to create value from Big Data. Using new technologies on existing data can provide additional dimensions of that data. Moreover, new types of analytics can help organizations to provide an improved overall customer experience, implement analytics in real time, and yield competitive intelligence about other offers (Capgemini, 2013e). Primary drivers for applying Big Data are efficiency, innovation, and compliance. It is important to find the right tool for a given workload and relational databases as used with Business Intelligence do not provide the most efficient way to store large unstructured data sets. Additionally, with Big Data it is possible to retrieve large amounts of data relatively quickly which makes it easier to comply with rules and regulations. Accordingly, it is possible to provide information instantly without preparing reports (De Vries, 2013b). Moreover, Big Data provides organizations with opportunities that were not feasible before because of the lacking technology or high expenses (Bakshi et al., 2013; Lopez, 2012).

Big Data gets increasing attention in the financial services industry because accurate, consistent management of both financial data and customer information is essential to be successful in this industry (Nasar & Bomers, 2012). It is striking that the volume of the data is often not the problem for organizations in the financial services industry. Organizations are used to work with large amounts of data, but the growing velocity and variety are more difficult to cope with. The variety of data can have an impact on organization’s risk measurements and its trading and investment performance. The difficulty lies in integrating the information from different sources by combining structured and unstructured data (Brown, 2012; Bean, 2013; Capgemini, 2013b; Capgemini, 2013e). The velocity of data is most important because it represents to what extent an organization can filter relevant information from the other two V’s as fast as possible to develop new products, to optimize processes, and to make better decisions (Capgemini, 2013e; De Vries, 2013b).

According to Lopez (2012), using Big Data can translate into operational, financial, and business gains, including one of the ultimate benefits gained from using Big Data which is that an organization can have quicker access to cleaner, more relevant data to drive insights and optimize decision making. Capgemini (2013e) adds to this that the use of Big Data allows organizations to analyze a much broader set of data about aspects of the organization as an integrated whole. Combining data from inside and outside the organization provides a more holistic view on different aspects in the organization such as customers, supply chains, products,
and processes. Capgemini (2013b) indicate these benefits as a more complete understanding of market conditions and evolving business trends, more accurate responses to customer needs, consistency of decision making, better risk management, and a competitive advantage. LaValle et al. (2011) and Database Trends and Applications (2012) add to this that the real benefit of Big Data is that it provides an opportunity to businesses by enabling users to run analytics and determine and predict market shifts, customer preferences, and product innovations. These benefits are also applicable in the financial services industry, however there is also another benefit from using Big Data analytics concerning regulatory compliance. Using Big Data to analyze the current regulatory environment can help organizations comply with today’s increasingly complex and everchanging regulatory requirements (Nasar & Bomers, 2012). However, this is difficult in the financial services industry because the rules and regulations often change faster than the IT landscape (De Vries, 2013b). Capgemini (2013e) summarizes these benefits in three parts: improving interaction with the ecosystem that results in better targeted marketing, individual service offerings and customer retention, improving business processes to understand the data more in detail and better predict future activity, and risk mitigation to manage risk compliance which is especially important in the financial services industry.

To gain from these benefits Big Data can provide, organizations need a solid plan. Organizations need to identify their key business drivers, how to put the processes into place, and need to make sure that they have the organizational alignment and skills to make it happen. In short, to be successful an organization should have the right people with the right skills on the right place and the right organizational structure to use Big Data (Bean, 2013).

Besides the benefits Big Data can provide, there are also challenges to the new development. One of the challenges is to integrate all information from different data silos together without separating it again in different silos. As said in Capgemini (2013b) and in De Vries (2013b), the biggest obstruction to use Big Data effectively for decision-making is the organizational silos. With the movement of Business Process Reengineering it was tried to eradicate the functional, vertical silos. Though, after a while new horizontal, process related silos were created. Additionally, today’s relational database architectures need to be modernized to adapt to the increasing volume, variety, and velocity and without removing the existing systems (Database Trends and Applications, 2012; DBTA, 2013a). Nasar and Bomers (2012) add that designing, developing, and implementing data management systems that live up to the access and speed requirements is a highly complicated process. Complexities in the process include the sheer number of stakeholders that are involved into the process, the multiple sources of data that should be integrated, the quality of the data, the integration of legacy systems with new systems, the short term operational challenges that often get priority above a long term project like this, and last the implementation problems such as cost overruns and schedule delays. Only buying and implementing a Big Data solution is not always the right answer to these challenges. Complementary initiatives should be implemented to be robust, such as digitalizing all the
information, identifying all the counterparties of the transactions, and governing the data by combining people, processes, and technology. Only when people, processes, and technology are all integrated it is possible to make the Big Data solution successful (Capgemini, 2013b; Capgemini, 2013e; De Vries, 2013b; Nasar & Bomers, 2012). This is also shown by LaValle et al. (2011), who indicate that many organizations have more trouble with the managerial and cultural barriers rather than data and technology barriers. To use analytic-driven insights, there should be a close connection to the business strategy, it should be easy to understand for the end-users, and it should be embedded into the processes, which are all difficult to implement. Brown et al. (2011), Capgemini (2013b), and De Vries (2013b) agree with this and indicate that one of the major complications for successfully implementing Big Data is talent. Organizations need people with deep analytical skills and managers with a clear understanding of how Big Data can be applied. So, additional substantial investments need to be made in recruitment and training. Moreover, time and money should be invested in change management that, amongst others, provides training to all employees to deal with the new business environment. The culture of the company should be changed to a culture in which the use of Big Data is more valued and rewarded (Capgemini, 2013b; De Vries, 2013b). This is only possible when every part of the organization is involved in the process and collaborates with each other (Capgemini, 2013e). However, cases that could be easily solved should not become more complex than they initially were (De Vries, 2013b).

Boyd and Crawford (2012) indicate challenges that are more general of nature. First of all, Big Data changes the definition of knowledge, which means that it creates a shift in how research is performed. Especially the process of research changes because of Big Data since no key questions are needed before starting any analysis. Secondly, using Big Data appears to be objective and accurate but this is misleading. The data that is used will still be interpreted by persons and data sets are often collected manually. This makes the research more subjective and less accurate as it seems on first glance. Third, Big Data refers to the volume of the data, however more data does not necessarily mean better data. The quality of the data does not have a direct positive relationship with the volume of data, which indicates that Big Data does not have to be better. Capgemini (2013e) agrees that the quality of the data will be more dependent on mature data management, governance, and stewardship with the volumes get bigger. Additionally, as described by Boyd and Crawford (2012), data sets are often reduced from larger data sets to fit the mathematical model that is used to analyze the data. However, when these data sets are reduced, data loses its meaning because it is taken out of context. Context is difficult to interpret at scale and when two datasets are modeled in the same way it does not mean that they are equivalent. On the contrary, in De Vries (2013b) it is indicated that Big Data can provide context to a large data set by discovering new patterns. Moreover, another challenge is that limited access to Big Data creates new digital divides because some organizations have full access to large data sets and skilled people while others cannot perform any investigation. Though, this division is not new, but originate from the forces of the market. Last, the issue of
privacy is addressed. Because a lot of information is available by using Big Data, this does not mean that it is ethical to use it all (Boyd & Crawford, 2012; Capgemini, 2013e; De Vries, 2013b). According to De Vries (2013b) there are little specific rules and regulations what is allowed concerning the privacy of individuals. Organizations often pay more attention to what a customer accepts as to what extent information can be used than what is allowed, because customers often accept less than what is prescribed by the law. Different parts of the world have different ideas about privacy and security, for example citizens of the Netherlands care more about their privacy than citizens of the United States. This is also confirmed by Brown et al. (2011) who indicate that greater access to personal information that is often demanded by Big Data put pressure on privacy and security issues. Therefore organizations have to be transparent about the use of the data of the customer. Moreover, it should be clear to whom the data belongs (De Vries, 2013b).

**Technology**

One of the main differences between Business Intelligence described above and Big Data is the difference in systems that are used. Whereas Business Intelligence retrieves information from Relational Database Management Systems (RDBMS), Big Data tries to combine data from different management systems together to provide more accurate information. According to Capgemini (2013c), before mastering Big Data, it should be realized that the analytical model is only as good as the information that is put in it, which means that sources need to be reliable. Moreover, the different data sets need common information to link them together. The mass of data (transactions and interactions) is created by different individual events which can be structured as a core of Parties, Objects and Locations. To do this, organizational governance is needed which consists of standards that define the structural format and definitions of information, and policies that define how standards will be enforced. Last, the individual transactional elements or events can be added and linked to the core Parties, Objects and Locations, to associate information in a consistent way across channels. This development of Big Data is also shown in figure 11.
Using distributed file systems that are available with Big Data technology allows organizations to get quick access to information. These Big Data technologies work best in cooperation with the original enterprise data warehouses as used with Business Intelligence (Nasar & Bomers, 2012). Existing infrastructures should not be dispensed, but current capabilities should be integrated with the new requirements of Big Data. This is also agreed upon by Capgemini (2013e) that indicates that Big Data is not only about technology but more about data management that combines traditional techniques with newer ones. The four steps that need to be taken to implement a Big Data solution are acquisition, marshalling, analysis, and action. The acquisition step includes collecting data from a variety of sources using extraction, transfer, and integration. The second step, marshalling, encompasses sorting and storing the data depending on the intended use. The analysis step contains finding insights and predictive modeling by combining information from the past with forward analytics. The last step, action, includes taking action on the insights to change business outcomes carried out by a human, a computer, or a combination of both. Especially this last step can be taken faster with the use of Big Data compared to the traditional approaches. The visualization of data becomes increasingly important for actions involving a human interaction. Because data sets are too large to make sense to a human, the visualization of data, including patterns and graphs, is used to base decisions on (Capgemini, 2013e).

The current RDBMS are not capable of dealing with the new data types (especially the unstructured data) and do not provide a convenient way of storing data that is necessary for Big Data. Collecting and storing both structured and unstructured data should be possible without a
clear idea how an organization can use it (DBTA, 2013b). Conventional relational databases are often structured and work with a well-defined schema. Using these databases for analysis is difficult when data sets are constantly changing because they are designed to optimize repeated queries. A new technology that does not use these repeated queries is the Hadoop technology (Bakshi et al., 2013; Capgemini, 2013e). The technology that belongs to the core of the Big Data revolution is MapReduce which was developed by Google to distribute and execute the computations where the data is originally stored. Later, the technology became available as an open source version via Hadoop (Arnold, 2012). Hadoop divides information into different blocks and does not need a predefined schema. It is designed to deal with large data sets by distributing the data over many servers enabling reliable, scalable, and distributed computing on these clusters of data (Capgemini, 2013e). Hadoop looks for patterns in data and dealing with unpredictable data sets and therefore focuses on flexibility and experimentation. Once a file is imported and stored in the system it can be accessed when it is needed. However, Hadoop lacks in providing direct responses when many people query at the same time. Organizations are used to work with Business Intelligence tools which provide an instant response on predefined questions, but Hadoop cannot provide that instant response (Bakshi et al., 2013). However, using Hadoop means that the data engineering process can be skipped and therefore it is possible to answer critical business questions in a shorter amount of time (Bean, 2013).

Additionally, Hadoop in itself does not create significant value. Hadoop provides the possibility to manage and process large amounts of data, but it should be complemented with other tools and platforms to optimize the data management. Hadoop only addresses the storage and analysis steps excluding the acquisition of data and actions that can be taken based on the analysis part (Capgemini, 2013e). Moreover, using Hadoop looks like a less expensive option than adding nodes to database appliances, however hardware is still needed to cope with the growing data volumes and performance service level agreements which raise the capital and operational costs. Also the ease of use is a challenge with Hadoop. Certain actions require specific skills that are hard to find which makes the use of Hadoop more difficult for organizations (Lopez, 2012). Other programs that use similar technology are jStart, Vivisimo, Vetas, Aster Data Systems, Vertica, Netezza, and Greenplum (DBTA, 2013a; IBM, 2013).

**Business Information Management**

The term Business Information Management encompasses the management process of converging data into information to gain insights and foresights in business performance. Whereas Business Intelligence and Big Data mainly focus on the techniques to transform data into insights, Business Information Management mainly focuses on the management behind it. It combines people, processes, and technology to create value from intelligence. Business Information Management consists of four parts: Business Intelligence, Big Data, User Experience, and Data Management. These four parts should be combined to create business value. Business Intelligence and Big Data refer to the technology, Data Management includes the
processes, and User Experience addresses the people part of Business Information Management. In literature, Business Intelligence and Business Information Management are often used interchangeably, though there is a difference in the focus of the terms (Negash, 2004; De Vries, 2013b).

According to Edgar (2005) Business Information Management provides a holistic understanding of management of information and information systems, the processes within a business, and the processes that transform data and information into knowledge. Business Information Management has the characteristic that it combines two perspectives: technological and business. The technological view includes the role of IT in the complete process; how IT can be used for gathering data, integrating data, analysis of data, and the distribution of information over the organization. This technological view refers more to the parts of Business Intelligence and Big Data. The business view includes the systematic process in which the core business strategy is the base for the business information needs that require data collection and analysis using knowledge and information. This knowledge and information then contribute to the business strategy. Capgemini (2013b) supports this and indicates that data can either help a manager in making a decision or it can automate decision-making. This business view refers more to the parts of Data Management and User Experience.

Business Information management is based on two assumptions. First of all, information is seen as a product that represents a certain value. Secondly, nowadays the success of an organization is determined by the extent to which organizations provide the possibility to employees to stand out by using information better and smarter (De Vries, 2013b). This can be used to deliver products and services that fulfill the needs of the customer.

**Business value of Business Intelligence**

To discover the business value of Business Intelligence it first should be clear what business value means in an IT setting. Only when this is clear it is possible to determine the business value of Business Intelligence. Therefore next paragraph will provide information on IT business value. Additionally the business value of Business Intelligence is explained.

**IT business value**

As there are multiple ways to define big Data, there are also multiple ways to define business value. In De Vries (2013b) business value reflects the performance that is visible and indicates what is actually performed and not what is said to be performed. In De Vries (2013b) business value is also described as everything that adds value to the business. Moreover it is described as the extent to which an organization is relevant to the customers in such a way that it is most efficient and effective on an emotional and rational level. More related to the financial services industry, in De Vries (2013b) business value is also described as having insights at the services that are provided and the acceptation of the customer.
A short definition comes from Masli et al. (2011) who define business value as “the organizational performance impacts of IT” (p. 83). According to Hitt and Brynjolfsson (1996), and Devaraj and Kohli (2003) these organizational performance impacts of IT include productivity enhancement, cost reduction, inventory reduction, profitability improvement, competitive advantage, and other measures of performance. However, according to Melville et al. (2004) the term performance is used for both intermediate process-level measures and organizational measures. Additionally they discovered that two formulations of performance are used in literature, efficiency-based performance and effectiveness-based performance. Efficiency emphasizes an internal perspective using metrics such as cost reduction and productivity, and shows the actual output compared to the input. Effectiveness emphasizes the achievement of organizational objectives in relation to the external environment by comparing the expected and actual output. To be complete, Melville et al. (2004) define IT business value as “the organizational performance impacts of IT at both the intermediate process level and the organization wide level, and comprising both efficiency impacts and competitive impacts” (p. 287).

Hitt & Brynjolfsson (1996) also showed that IT business value can be defined in multiple ways. This paper indicated that the question if the IT investment generates business value is not just one simple question, but contains of multiple questions about productivity, business profitability, and value for consumers. The productivity reflects the efficiency of Melville et al. (2004) and indicates the marginal benefits of IT. Business profitability reflects the effectiveness of Melville et al (2004) and dictates if the benefits that are created by IT can be used to gain competitive advantage. The value for consumers reflects the magnitude of the benefits. These three areas are distinct and have different implications for how managers and researchers see the investment.

Besides the different ways to define business value of an IT investment, there are more issues that have an influence on this concept. According to Ahituv (1980), it is important to understand whose value it is that is talked about. The value can be individual, team-based, organizational or any other group of individuals. Additionally it is important to understand about what type of value is talked about. Generally there are three types of value: the perceived value, the revealed or realistic value, and the normative value (Mock, 1971; Ahituv, 1980). The perceived value is the value that is discerned by the user and indicates the subjective evaluation of the information’s worth. It is only available when experiments and interviews are performed (Munro & Davis, 1977; King & Rodriguez, 1978). The revealed value is the marginal improvement of the performance of the user after receiving the information. This value reflects the most realistic impact of the information. (Ahituv, 1980). The third type of value, normative value, is the difference between the expected payoff when information is optimally used and the expected payoff when no information is given. The three types of value are distinct and there is no ultimate answer which value should be considered (Ahituv, 1980).
Corporate Performance Management

Determining business value is only possible when performance is measured. The part of an organization that deals with how business value can be created is Corporate Performance Management (CPM). CPM, also known as Enterprise Performance Management (EPM) or Business Performance Management (BPM), entails the continuous process in which methodologies, metrics, processes, and systems are used to manage and monitor performance by achieving goals that induce the ultimate goal of realizing the defined strategy (Buytendijk, 2005; Kemper et al., 2013). CPM functions as a communication device between the strategic level of management and the business operations on a tactical and operational level. Strategic objectives and performance targets are top-down communicated while reliable and consistent information is communicated bottom-up to feed the steering and control process (Van Wunnik et al., 2006). The CPM process contains four steps. First of all, norms should be set for every measurement. Second, the actual measurements should be done. Third, the actual measurements have to be compared to the norms set during the first step to discover potential deviation. The last step includes taking a certain action that is based on the conclusions from step 3 (Buytendijk, 2005).

Performance is often expressed by financial indicators, because the business decisions resulting from the performance outcomes are mostly based on financial measures. Moreover, financial indicators are more tangible than the non-financial ones which makes it easier to understand. However, focusing on financial indicators has some disadvantages. First of all, financial measures only indicate something about the organization itself, not about the competition. Additionally, financial indicators are based on historical data and do not provide information for the future. Last, using financial indicators is aimed at the shareholder, forgetting other stakeholders such as customers and employees.

These disadvantages are taken into account at different frameworks that could be used to manage performance. The research work of De Vries (2013a) shows many different performance frameworks. One of the frameworks that is widely known and used in practice is the Balanced Scorecard of Kaplan and Norton (1996; 2000). The Balanced Scorecard approach allows a comprehensive view of the nature of the IT investment. Moreover, it highlights how IT investments influence the processes and the overall organizational performance in one framework (Masli et al., 2011). According to Kaplan and Norton (1996) the traditional accounting models that could be used to measure the business value of IT do not incorporate the intangible assets and company capabilities such as high quality products and services, skilled employees, and satisfied customers that are critical for success. Therefore a different framework is developed that includes both the historical-cost financial accounting model and new long-range competitive capabilities, which is the Balanced Scorecard. It complements the financial measures of past performance with the measures of the drivers of future performance. The objectives and measures are derived from an organization’s vision and strategy. Organizational performance is viewed from four perspectives: financial, internal business process, customer, and learning and growth (Kaplan & Norton, 1996). The financial perspective is valuable to
summarize the readily measurable economic consequences of actions that are already taken. The financial performance measures show if a company’s strategy, implementation, and execution contribute to bottom-line improvement. The internal business process perspective is the perspective in which executives can identify the critical internal processes that enable to deliver the value propositions that attract and retain customers in targeted market segments, and satisfy shareholder expectations of financial returns. It includes both the short term operations cycle as well as the long term innovation cycle. The customer perspective identifies customer and market segments in which will be competed and how the performance can be measured in these targeted segments. It includes generic measures of successful outcomes and specific measures of the value propositions that will be delivered. The learning and growth perspective “identifies the infrastructure that the organization must build to create long-term growth and improvement” (Kaplan & Norton, 1996, p. 28). The organizational learning and growth comes from the sources people, systems, and organizational procedures. The other perspectives of the Balanced Scorecard identify gaps between the capabilities of people, the capabilities of systems, and the capabilities of procedures that are needed to achieve performance. This perspective incorporates the investments that organizations have to do to close these gaps. All perspectives are divided in objectives, measures, targets, and initiatives. The objectives indicate what an organization wants to evaluate. The measures are used to realize the objectives. The targets give an indication of the current and expected value for measures of objectives that have to be realized. Lastly, the initiatives are assigned for each objective and the responsibility is determined (Kaplan & Norton, 1996).

Many organizations already use financial and non-financial measures for organizational performance, however the non-financial measures are often only used to improve locally at the front-line and at customer-facing operations. Financial measures are often used to only control short-term operations. The Balanced Scorecard emphasizes that both type of measures should be part of the information system for all employees at all levels of the organization. The measures are deduced from a top-down process driven by the vision and strategy of the organization and represent a balance between the external measures for shareholders and customers, and the internal measures. The Balanced Scorecard is more than a collection of critical success factors, because it consists of a linked series of objectives and measures that are consistent and mutually reinforcing. The scorecard incorporates a complex set of cause and effect relationships among the critical variables that describe the trajectory of the strategy. The financial results are driven by customer satisfaction which is driven by internal processes. Underneath these layers, there is the foundation of the learning and growth perspective (Kaplan & Norton, 2000). Moreover, it also provides a mix between outcome measures and performance drivers such as cycle time and defect rates. Furthermore, a feedback loop is included which makes sure that the objectives that are determined are questioned once in a while if they are still consistent with current observations and experiences (Kaplan & Norton, 1996).
Researchers are not all positive about the Balanced Scorecard. The paper of Nørreklit (2000) shows some critique on the use of the Balanced Scorecard. Suggestions are proposed to tackle these negative points. One of the suggestions is to see the relationship between the perspectives not as causal but as coherence between measurements. Coherence focuses on whether the measurements match or complement each other. To construct coherent performance measurements, resources and strategy should be developed on an approximately coherent way. It should not be forgotten to have a strategic dialogue when measurements are formulated to ensure that the actors in individual activities have commitment.

*(Business) value of Business Intelligence*

Also for the Business Intelligence business value contains everything that adds value to the company. In literature there is no unanimity about how to measure the business value of Business Intelligence.

The paper of Lönnqvist and Pirttimäki (2006) indicates that it is difficult to get unanimity about how to measure the business value of Business Intelligence investments. First of all it should be clear what the purpose of the measurement is before it can be decided what and how to measure it. Different purposes for measuring performance are decision making, control, guidance, external communication, and education and learning (Simons, 2000). The main purposes for measuring the performance of Business Intelligence are to prove that the investment is worth it and to manage the process to ensure that users are satisfied and the process is efficient. Therefore it is necessary that it is clear how much it costs to apply Business Intelligence and what the benefits are (Lönnqvist & Pirttimäki, 2006).

To calculate the total costs, labor costs, information purchases, and other expenses should be included (Lönnqvist & Pirttimäki, 2006). The method Total Cost of Ownership proposed by Degraeve et al. (2004) could be used to identify all relevant costs. However, according to Lönnqvist and Pirttimäki (2006), measuring the benefits is more difficult because benefits of the investments are often non-financial and even intangible, such as improved quality and timeliness of information. Although these benefits should lead to financial outcomes, it is possible that there is a time lag in this relationship and therefore it is difficult to measure it. Additionally, it is even difficult to assess the financial benefits. The outcome of Business Intelligence is intelligence which is some kind of processed information. Determining the value of this information in financial terms is difficult. Therefore methods such as ROI and net present value are problematic. Though, Davison (2001) developed a measurement model that can be used to calculate the return on an investment by assessing the value of individual Business Intelligence projects and measuring the output by assessing the effects. The incremental value that is added to a decision could be presented in financial terms and the input is the cost of carrying out the project. Together, the return on investment is (outputs – inputs)/inputs. According to Lönnqvist and Pirttimäki (2006), the calculation can be unreliable because the outputs are based on qualitative assessments. Therefore it is better to measure the effectiveness. Four measures are
identified: time savings, cost savings, cost avoidance, and revenue enhancement. The effectiveness can be measured by evaluating the contribution of Business Intelligence to a specific decision or action and how it benefits the organization. Moreover, it is possible to measure the effectiveness subjectively by evaluating the perceived customer satisfaction. One thing that should not be forgotten is that measuring the performance should also be done in the context of managing the process. The goal is to efficiently produce valuable intelligence for the specific needs of the users.

The paper of Williams and Williams (2003) also describes how different methods can be used to measure the performance. According to Williams and Williams (2003), often the business value of an investment is measured as the net present value of the after-tax cash flows. Therefore investments in Business Intelligence should be subjected to an assessment of how the investment results in increased revenues and/or reduced costs. All benefits that are proposed such as agility, responsiveness, information sharing, and flexibility are meaningless when these attributes are not defined in operational terms and realized through business processes to influence cash flows. That is why the business value of Business Intelligence can be defined as how an organization uses Business Intelligence to improve management processes such as planning and monitoring to change the cash flows from a management side, and to improve operational processes such as sales campaign execution and purchasing to increase revenue and/or reduce costs from a business side. This means that to capture the business value, organizations should go beyond the technical implementation of the systems. Preconditions that should be met to capture the business value are strategic alignment, process engineering, change management, BI technical development, and BI project management. Only when these conditions are met it is possible to create business value for the specific environment. The business value does not only contain a traditional return on investment analysis which estimates the net present value of after-tax cash flows, but it should be taken from a broader perspective including opportunity analysis, readiness assessment, process engineering, and change analysis. The opportunity analysis combines environmental, industry, and business strategy analysis with an assessment of the use of Business Intelligence to enable critical strategies and support key business processes. The readiness assessment defines the organizational, business, and technical readiness to deliver information. Moreover it can be extended with a maturity assessment to evaluate organizational management. Process engineering also influences the business value by determining how Business Intelligence exactly will be used in the context of the management processes to improve those that drive the cash flow. Last, a change analysis is also performed to extend the results of process engineering by indicating which level of process and individual change is required to create business value (Williams & Williams, 2003).

Elbashir et al. (2008) agree with Lönnqvist and Pirttimäki (2006) that performance measures need to be developed specific for Business Intelligence that capture the context of use in terms of management objectives as well as specific capabilities. To achieve the benefits of investing in Business Intelligence such as improved decision making, the systems need to be effectively
integrated into management and operational processes (Williams & Williams, 2003; De Voe & Neal, 2005; Elbashir et al., 2008). Therefore performance measurement impacts two levels by improving the efficiency and effectiveness of the organizational structure and business processes, and by outperforming other organizations in the industry. Hence, the business value is also defined for two levels. The business process performance includes the operational efficiency and effectiveness improvements in business processes that are empowered by Business Intelligence systems. These improvements can lead to business value, but that is dependent on other factors such as the scope of the business process, and the competitive environment. The other level accumulates the enabled performance across the organization. The study of Elbashir et al. (2008) shows that customer intelligence, supplier relations, and internal efficiency all have an influence on business process performance and eventually business process performance effects organizational performance. The last relationship is stronger for non-service industries compared to service industries such as the financial services industry.

According to Lin et al. (2009), not only efficiency and effectiveness determine the business value of Business Intelligence, but it is also important that the systems can be accepted by the users, that it satisfies the needs of the users, and that the objectives that are composed are met after implementation. This can be measured by contents correctness, easiness of the operation, resilience of the format, integrity of the output, real-time nature, integration and safety of the system, and credibility of the output.

Ghazanfari et al. (2011) take a different approach for evaluating the business value of Business Intelligence. The paper proposes an evaluation tool that could be used to determine the level of intelligence of the Business Intelligence competencies. The six factors that were discovered are analytical and intelligent decision-support, providing related experiment and integration with environmental information, optimization and recommended model, reasoning, enhanced decision-making tools, and stakeholders’ satisfaction. The intelligence of the organizational systems is evaluated by 34 criteria divided between the six factors. This results in a dashboard with the level of intelligence per factor, together indicating the level of intelligence and which parts should be improved to create more business value.

Although above measurements could be used to assess the business value of Business Intelligence, more comprehensive methods exist to have a better overview of all possible measures that influence the organizational performance. Literature proposes different balanced performance measurement frameworks that can be used to assess the business value of Business Intelligence, such as the Balanced Scorecard of Kaplan and Norton (1996; 2000). Lönnqvist and Pirttimäki (2006) show that the main principles are the same in different frameworks divided over five stages. First performance measures are chosen based on an organization’s strategy. Next, the success factors are chosen from the different perspectives included in the framework. In the third stage measurement is limited to a number of critical success factors. After that the measurement system is designed in such a way that causal relationships between the success factors are shown. During the last stage the measurement system can be used as a tool to
implement and communicate the initial strategy. The success factors can be directly or indirectly measured. A direct measurement shows the effect of Business Intelligence while an indirect measurement shows the utilization of the intelligence. Moreover, the data that is used to indicate a success factor could be objective or subjective. Although there are different performance frameworks that could be used to assess the business value, no framework is specifically developed for Business Intelligence. Additionally, different types of measurement are discussed, such as financial and non-financial measurements, but there is no agreement on the best way to assess organizational performance.
Figure 12: First draft of conceptual framework
Figure 13: Second draft of conceptual framework
Appendix E: Results questionnaire

E1: Little’s MCAR $\chi^2$ test

a. Little’s MCAR test: Chi-Square = 837,524, DF = 1932, Sig. = 1,000
b. The EM algorithm failed to converge in 25 iterations.

E2: Descriptive statistics dataset

Table 6: Descriptive statistics dataset

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>crossselling retention</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,423</td>
</tr>
<tr>
<td>customizeddeals satisfaction</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,807</td>
</tr>
<tr>
<td>packages_synchronizingIT</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,461</td>
</tr>
<tr>
<td>packages_legacynew</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,346</td>
</tr>
<tr>
<td>projectteam_training_errors</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,500</td>
</tr>
<tr>
<td>newemployees_training_errors</td>
<td>27</td>
<td>1,00</td>
<td>4,00</td>
<td>3,423</td>
</tr>
<tr>
<td>employeesatisfaction_employeecosts</td>
<td>27</td>
<td>1,00</td>
<td>4,00</td>
<td>3,000</td>
</tr>
<tr>
<td>training_employeecosts</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,280</td>
</tr>
<tr>
<td>projectteam_employeesatisfaction</td>
<td>27</td>
<td>1,00</td>
<td>4,00</td>
<td>3,200</td>
</tr>
<tr>
<td>newemployees_employeesatisfaction</td>
<td>27</td>
<td>1,00</td>
<td>4,00</td>
<td>3,080</td>
</tr>
<tr>
<td>training_errors_investments</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,269</td>
</tr>
<tr>
<td>projectteam_training</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,346</td>
</tr>
<tr>
<td>Gender</td>
<td>27</td>
<td>1</td>
<td>2</td>
<td>1,22</td>
</tr>
<tr>
<td>Age</td>
<td>27</td>
<td>2</td>
<td>4</td>
<td>2,70</td>
</tr>
<tr>
<td>Nationality</td>
<td>27</td>
<td>1</td>
<td>4</td>
<td>1,33</td>
</tr>
<tr>
<td>Education</td>
<td>27</td>
<td>1</td>
<td>4</td>
<td>3,33</td>
</tr>
<tr>
<td>Current_sector</td>
<td>27</td>
<td>1</td>
<td>4</td>
<td>1,59</td>
</tr>
<tr>
<td>Current_company</td>
<td>27</td>
<td>1</td>
<td>2</td>
<td>1,22</td>
</tr>
<tr>
<td>job_status</td>
<td>27</td>
<td>1</td>
<td>3</td>
<td>2,67</td>
</tr>
<tr>
<td>number_years_worked_current_organization</td>
<td>27</td>
<td>1</td>
<td>5</td>
<td>3,30</td>
</tr>
<tr>
<td>bigdata_increase_revenue</td>
<td>27</td>
<td>2</td>
<td>5</td>
<td>3,96</td>
</tr>
<tr>
<td>bigdata_decrease_costs</td>
<td>27</td>
<td>2,00</td>
<td>5,00</td>
<td>3,556</td>
</tr>
<tr>
<td>sales_revenue</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,592</td>
</tr>
<tr>
<td>retention_revenue</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,703</td>
</tr>
<tr>
<td>satisfaction_revenue</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,741</td>
</tr>
<tr>
<td>marketshare_revenue</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,370</td>
</tr>
<tr>
<td>crossselling_sales</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,629</td>
</tr>
<tr>
<td>customizeddeals_sales</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,778</td>
</tr>
</tbody>
</table>
Table 6: Descriptive statistics dataset (continued)

<table>
<thead>
<tr>
<th>Description</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>customizeddeals_retention</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.8148</td>
</tr>
<tr>
<td>insights_competing_market_marketshare</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6296</td>
</tr>
<tr>
<td>payments_crossselling</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6296</td>
</tr>
<tr>
<td>savings_crossselling</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5556</td>
</tr>
<tr>
<td>financing_crossselling</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.4444</td>
</tr>
<tr>
<td>investments_crossselling</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6667</td>
</tr>
<tr>
<td>insurances_crossselling</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.4815</td>
</tr>
<tr>
<td>payments_customizeddeals</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5185</td>
</tr>
<tr>
<td>savings_customizeddeals</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5926</td>
</tr>
<tr>
<td>financing_customizeddeals</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.3704</td>
</tr>
<tr>
<td>investments_customizeddeals</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5926</td>
</tr>
<tr>
<td>insurances_customizeddeals</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.4815</td>
</tr>
<tr>
<td>payments_competingmarket</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6667</td>
</tr>
<tr>
<td>savings_competingmarket</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5185</td>
</tr>
<tr>
<td>financing_competingmarket</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.4074</td>
</tr>
<tr>
<td>investments_competingmarket</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6296</td>
</tr>
<tr>
<td>insurances_competingmarket</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.4444</td>
</tr>
<tr>
<td>employeecosts_costs</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>2.9259</td>
</tr>
<tr>
<td>creditscoring_costs</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5185</td>
</tr>
<tr>
<td>fraud_costs</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>4.0370</td>
</tr>
<tr>
<td>payments_creditscoring</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.7407</td>
</tr>
<tr>
<td>savings_creditscoring</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6296</td>
</tr>
<tr>
<td>financing_creditscoring</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5185</td>
</tr>
<tr>
<td>investments_creditscoring</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6296</td>
</tr>
<tr>
<td>insurances_creditscoring</td>
<td>27</td>
<td>1.00</td>
<td>4.00</td>
<td>3.3704</td>
</tr>
<tr>
<td>payments_fraud</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.9259</td>
</tr>
<tr>
<td>savings_fraud</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5926</td>
</tr>
<tr>
<td>financing_fraud</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.4444</td>
</tr>
<tr>
<td>investments_fraud</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6667</td>
</tr>
<tr>
<td>insurances_fraud</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.3333</td>
</tr>
<tr>
<td>synchronizingIT_payments</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6667</td>
</tr>
<tr>
<td>legacynew_payments</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.6296</td>
</tr>
<tr>
<td>visualization_analytics_payments</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.8889</td>
</tr>
<tr>
<td>training_errors_payments</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.2222</td>
</tr>
<tr>
<td>synchronizingIT_savings</td>
<td>27</td>
<td>1.00</td>
<td>5.00</td>
<td>3.5926</td>
</tr>
</tbody>
</table>
Table 6: Descriptive statistics dataset (continued)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>legacynew_savings</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,5926</td>
</tr>
<tr>
<td>visualization_analytics_savings</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,8519</td>
</tr>
<tr>
<td>training_errors_savings</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,2222</td>
</tr>
<tr>
<td>synchronizingIT_financing</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,6667</td>
</tr>
<tr>
<td>legacynew_financing</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,6667</td>
</tr>
<tr>
<td>visualization_analytics_financing</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,2593</td>
</tr>
<tr>
<td>training_errors_financing</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,2593</td>
</tr>
<tr>
<td>synchronizingIT_investments</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,6296</td>
</tr>
<tr>
<td>legacynew_investments</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,6296</td>
</tr>
<tr>
<td>visualization_analytics_investments</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,9259</td>
</tr>
<tr>
<td>training_errors_insurances</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,2593</td>
</tr>
<tr>
<td>synchronizingIT_insurances</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,6296</td>
</tr>
<tr>
<td>legacynew_insurances</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,9259</td>
</tr>
<tr>
<td>visualization_analytics_insurances</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,6296</td>
</tr>
<tr>
<td>training_errors_insurances</td>
<td>27</td>
<td>1,00</td>
<td>5,00</td>
<td>3,2222</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E3: Demographics

Table 7: Descriptives gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>21</td>
<td>77,8</td>
<td>77,8</td>
<td>77,8</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>22,2</td>
<td>22,2</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Descriptives age

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-35</td>
<td>12</td>
<td>44,4</td>
<td>44,4</td>
<td>44,4</td>
</tr>
<tr>
<td>36-50</td>
<td>11</td>
<td>40,7</td>
<td>40,7</td>
<td>85,2</td>
</tr>
<tr>
<td>51-65</td>
<td>4</td>
<td>14,8</td>
<td>14,8</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>
### Table 9: Descriptives nationality

<table>
<thead>
<tr>
<th>Nationality</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch</td>
<td>20</td>
<td>74,1</td>
<td>74,1</td>
<td>74,1</td>
</tr>
<tr>
<td>Indian</td>
<td>6</td>
<td>22,2</td>
<td>22,2</td>
<td>96,3</td>
</tr>
<tr>
<td>Norwegian</td>
<td>1</td>
<td>3,7</td>
<td>3,7</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>

### Table 10: Descriptives education

<table>
<thead>
<tr>
<th>Education</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>high school</td>
<td>3</td>
<td>11,1</td>
<td>11,1</td>
<td>11,1</td>
</tr>
<tr>
<td>bachelor degree</td>
<td>9</td>
<td>33,3</td>
<td>33,3</td>
<td>44,4</td>
</tr>
<tr>
<td>master degree</td>
<td>15</td>
<td>55,6</td>
<td>55,6</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>

### Table 11: Descriptives current sector

<table>
<thead>
<tr>
<th>Current sector</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial services</td>
<td>20</td>
<td>74,1</td>
<td>74,1</td>
<td>74,1</td>
</tr>
<tr>
<td>Retail</td>
<td>2</td>
<td>7,4</td>
<td>7,4</td>
<td>81,5</td>
</tr>
<tr>
<td>Care</td>
<td>1</td>
<td>3,7</td>
<td>3,7</td>
<td>85,2</td>
</tr>
<tr>
<td>IT consulting</td>
<td>4</td>
<td>14,8</td>
<td>14,8</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>
Table 12: Descriptives current company

<table>
<thead>
<tr>
<th>Current company</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capgemini</td>
<td>21</td>
<td>77,8</td>
<td>77,8</td>
<td>77,8</td>
</tr>
<tr>
<td>Valid Other</td>
<td>6</td>
<td>22,2</td>
<td>22,2</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>

Table 13: Descriptives job status

<table>
<thead>
<tr>
<th>Job status</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>first level supervisor</td>
<td>2</td>
<td>7,4</td>
<td>7,4</td>
<td>7,4</td>
</tr>
<tr>
<td>middle management</td>
<td>5</td>
<td>18,5</td>
<td>18,5</td>
<td>25,9</td>
</tr>
<tr>
<td>Nonmanagerial</td>
<td>20</td>
<td>74,1</td>
<td>74,1</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>

Table 14: Descriptives number of years worked in current organization

<table>
<thead>
<tr>
<th>Number of years worked in current organization</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than 1</td>
<td>1</td>
<td>3,8</td>
<td>3,8</td>
<td>3,8</td>
</tr>
<tr>
<td>1-2</td>
<td>10</td>
<td>37,0</td>
<td>37,0</td>
<td>40,8</td>
</tr>
<tr>
<td>3-5</td>
<td>4</td>
<td>14,8</td>
<td>14,8</td>
<td>55,6</td>
</tr>
<tr>
<td>6-10</td>
<td>4</td>
<td>14,8</td>
<td>14,8</td>
<td>70,4</td>
</tr>
<tr>
<td>over 10</td>
<td>8</td>
<td>29,6</td>
<td>29,6</td>
<td>100,0</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>100,0</td>
<td>100,0</td>
<td></td>
</tr>
</tbody>
</table>

E4: Cronbach’s Alpha
The table below shows the Cronbach’s Alpha (α) when a variable would be deleted. The result shows that it is not necessary to delete a variable to increase the reliability of the data.
Table 15: Cronbach’s Alpha for separate variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Corrected Item-Total Correlation</th>
<th>Cronbach’s Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>bigdata_increase_revenue</td>
<td>241.1000</td>
<td>1758.347</td>
<td>.511</td>
<td>.983</td>
</tr>
<tr>
<td>bigdata_decrease_costs</td>
<td>241.5074</td>
<td>1750.729</td>
<td>.459</td>
<td>.983</td>
</tr>
<tr>
<td>sales_revenue</td>
<td>241.4704</td>
<td>1739.918</td>
<td>.681</td>
<td>.983</td>
</tr>
<tr>
<td>retention_revenue</td>
<td>241.3593</td>
<td>1742.025</td>
<td>.611</td>
<td>.983</td>
</tr>
<tr>
<td>satisfaction_revenue</td>
<td>241.3222</td>
<td>1737.767</td>
<td>.732</td>
<td>.982</td>
</tr>
<tr>
<td>marketshare_revenue</td>
<td>241.6926</td>
<td>1750.594</td>
<td>.640</td>
<td>.983</td>
</tr>
<tr>
<td>crossselling_sales</td>
<td>241.4333</td>
<td>1745.523</td>
<td>.676</td>
<td>.983</td>
</tr>
<tr>
<td>customizeddeals_sales</td>
<td>241.2852</td>
<td>1742.893</td>
<td>.612</td>
<td>.983</td>
</tr>
<tr>
<td>crossselling_retention</td>
<td>241.6400</td>
<td>1746.872</td>
<td>.656</td>
<td>.983</td>
</tr>
<tr>
<td>customizeddeals_retention</td>
<td>241.2481</td>
<td>1737.182</td>
<td>.724</td>
<td>.982</td>
</tr>
<tr>
<td>customizeddeals_satisfaction</td>
<td>241.2552</td>
<td>1732.983</td>
<td>.746</td>
<td>.982</td>
</tr>
<tr>
<td>insights_competing_market_marketshare</td>
<td>241.4333</td>
<td>1752.146</td>
<td>.550</td>
<td>.983</td>
</tr>
<tr>
<td>payments_crossselling</td>
<td>241.4333</td>
<td>1737.330</td>
<td>.754</td>
<td>.982</td>
</tr>
<tr>
<td>savings_crossselling</td>
<td>241.5074</td>
<td>1730.978</td>
<td>.795</td>
<td>.982</td>
</tr>
<tr>
<td>financing_crossselling</td>
<td>241.6185</td>
<td>1732.690</td>
<td>.740</td>
<td>.982</td>
</tr>
<tr>
<td>investments_crossselling</td>
<td>241.3963</td>
<td>1734.616</td>
<td>.797</td>
<td>.982</td>
</tr>
<tr>
<td>insurances_crossselling</td>
<td>241.5815</td>
<td>1731.467</td>
<td>.788</td>
<td>.982</td>
</tr>
<tr>
<td>payments_customizeddeals</td>
<td>241.5444</td>
<td>1740.486</td>
<td>.703</td>
<td>.983</td>
</tr>
<tr>
<td>savings_customizeddeals</td>
<td>241.4704</td>
<td>1727.747</td>
<td>.841</td>
<td>.982</td>
</tr>
<tr>
<td>financing_customizeddeals</td>
<td>241.6926</td>
<td>1735.638</td>
<td>.741</td>
<td>.982</td>
</tr>
<tr>
<td>investments_customizeddeals</td>
<td>241.4704</td>
<td>1732.396</td>
<td>.818</td>
<td>.982</td>
</tr>
<tr>
<td>insurances_customizeddeals</td>
<td>241.5815</td>
<td>1726.957</td>
<td>.811</td>
<td>.982</td>
</tr>
<tr>
<td>payments_competingmarket</td>
<td>241.3963</td>
<td>1749.477</td>
<td>.711</td>
<td>.983</td>
</tr>
<tr>
<td>savings_competingmarket</td>
<td>241.5444</td>
<td>1743.116</td>
<td>.703</td>
<td>.983</td>
</tr>
<tr>
<td>financing_competingmarket</td>
<td>241.6556</td>
<td>1748.092</td>
<td>.635</td>
<td>.983</td>
</tr>
<tr>
<td>investments_competingmarket</td>
<td>241.4333</td>
<td>1748.546</td>
<td>.718</td>
<td>.983</td>
</tr>
<tr>
<td>insurances_competingmarket</td>
<td>241.6185</td>
<td>1743.718</td>
<td>.695</td>
<td>.983</td>
</tr>
<tr>
<td>employeecosts_costs</td>
<td>242.1370</td>
<td>1744.457</td>
<td>.557</td>
<td>.983</td>
</tr>
<tr>
<td>creditscoring_costs</td>
<td>241.5444</td>
<td>1741.966</td>
<td>.576</td>
<td>.983</td>
</tr>
<tr>
<td>fraud_costs</td>
<td>241.0259</td>
<td>1738.867</td>
<td>.633</td>
<td>.983</td>
</tr>
<tr>
<td>payments_creditscoring</td>
<td>241.3222</td>
<td>1742.237</td>
<td>.671</td>
<td>.983</td>
</tr>
<tr>
<td>savings_creditscoring</td>
<td>241.4333</td>
<td>1737.539</td>
<td>.716</td>
<td>.983</td>
</tr>
<tr>
<td>financing_creditscoring</td>
<td>241.5444</td>
<td>1749.003</td>
<td>.587</td>
<td>.983</td>
</tr>
</tbody>
</table>
## Table 15: Cronbach’s Alpha for separate variables (continued)

<table>
<thead>
<tr>
<th></th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Corrected Item-Total Correlation</th>
<th>Cronbach’s Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>investments_creditscoring</td>
<td>241.4333</td>
<td>1750.496</td>
<td>.605</td>
<td>.983</td>
</tr>
<tr>
<td>insurances_creditscoring</td>
<td>241.6926</td>
<td>1752.044</td>
<td>.552</td>
<td>.983</td>
</tr>
<tr>
<td>payments_fraud</td>
<td>241.1370</td>
<td>1743.208</td>
<td>.648</td>
<td>.983</td>
</tr>
<tr>
<td>savings_fraud</td>
<td>241.4704</td>
<td>1743.264</td>
<td>.610</td>
<td>.983</td>
</tr>
<tr>
<td>financing_fraud</td>
<td>241.6185</td>
<td>1747.698</td>
<td>.606</td>
<td>.983</td>
</tr>
<tr>
<td>investments_fraud</td>
<td>241.3963</td>
<td>1749.488</td>
<td>.563</td>
<td>.983</td>
</tr>
<tr>
<td>insurances_fraud</td>
<td>241.7296</td>
<td>1749.965</td>
<td>.557</td>
<td>.983</td>
</tr>
<tr>
<td>synchronizingIT_payments</td>
<td>241.3963</td>
<td>1733.620</td>
<td>.811</td>
<td>.982</td>
</tr>
<tr>
<td>legacynew_payments</td>
<td>241.4333</td>
<td>1742.839</td>
<td>.618</td>
<td>.983</td>
</tr>
<tr>
<td>visualization_analytics_payments</td>
<td>241.1741</td>
<td>1738.587</td>
<td>.730</td>
<td>.982</td>
</tr>
<tr>
<td>training_errors_payments</td>
<td>241.8407</td>
<td>1746.755</td>
<td>.542</td>
<td>.983</td>
</tr>
<tr>
<td>synchronizingIT_savings</td>
<td>241.4704</td>
<td>1734.169</td>
<td>.836</td>
<td>.982</td>
</tr>
<tr>
<td>legacynew_savings</td>
<td>241.4704</td>
<td>1745.458</td>
<td>.608</td>
<td>.983</td>
</tr>
<tr>
<td>visualization_analytics_savings</td>
<td>241.2111</td>
<td>1739.765</td>
<td>.701</td>
<td>.983</td>
</tr>
<tr>
<td>training_errors_savings</td>
<td>241.8407</td>
<td>1746.755</td>
<td>.542</td>
<td>.983</td>
</tr>
<tr>
<td>synchronizingIT_financing</td>
<td>241.3963</td>
<td>1733.466</td>
<td>.859</td>
<td>.982</td>
</tr>
<tr>
<td>legacynew_financing</td>
<td>241.3963</td>
<td>1744.756</td>
<td>.625</td>
<td>.983</td>
</tr>
<tr>
<td>visualization_analytics_financing</td>
<td>241.1370</td>
<td>1737.330</td>
<td>.763</td>
<td>.982</td>
</tr>
<tr>
<td>training_errors_financing</td>
<td>241.8037</td>
<td>1745.549</td>
<td>.551</td>
<td>.983</td>
</tr>
<tr>
<td>synchronizingIT_investments</td>
<td>241.4333</td>
<td>1732.935</td>
<td>.859</td>
<td>.982</td>
</tr>
<tr>
<td>legacynew_investments</td>
<td>241.4333</td>
<td>1744.224</td>
<td>.628</td>
<td>.983</td>
</tr>
<tr>
<td>visualization_analytics_investments</td>
<td>241.1000</td>
<td>1736.147</td>
<td>.758</td>
<td>.982</td>
</tr>
<tr>
<td>training_errors_investments</td>
<td>241.7937</td>
<td>1744.669</td>
<td>.562</td>
<td>.983</td>
</tr>
<tr>
<td>synchronizingIT_insurances</td>
<td>241.4333</td>
<td>1732.935</td>
<td>.859</td>
<td>.982</td>
</tr>
<tr>
<td>legacynew_insurances</td>
<td>241.4333</td>
<td>1744.224</td>
<td>.628</td>
<td>.983</td>
</tr>
<tr>
<td>visualization_analytics_insurances</td>
<td>241.1370</td>
<td>1738.792</td>
<td>.742</td>
<td>.982</td>
</tr>
<tr>
<td>training_errors_insurances</td>
<td>241.8407</td>
<td>1744.909</td>
<td>.587</td>
<td>.983</td>
</tr>
<tr>
<td>packages_synchronizingIT</td>
<td>241.6015</td>
<td>1742.855</td>
<td>.643</td>
<td>.983</td>
</tr>
<tr>
<td>packages_legacynew</td>
<td>241.7167</td>
<td>1751.159</td>
<td>.570</td>
<td>.983</td>
</tr>
<tr>
<td>projectteam_training_errors</td>
<td>241.5630</td>
<td>1748.677</td>
<td>.627</td>
<td>.983</td>
</tr>
<tr>
<td>newemployees_training_errors</td>
<td>241.6400</td>
<td>1741.993</td>
<td>.824</td>
<td>.982</td>
</tr>
<tr>
<td>employeesatisfaction_employeecosts</td>
<td>242.0630</td>
<td>1739.536</td>
<td>.729</td>
<td>.982</td>
</tr>
</tbody>
</table>
Table 15: Cronbach’s Alpha for separate variables (continued)

<table>
<thead>
<tr>
<th>Scale Name</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Corrected Item-Total Correlation</th>
<th>Cronbach’s Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>training_employeecosts</td>
<td>241,7830</td>
<td>1745,722</td>
<td>.699</td>
<td>.983</td>
</tr>
<tr>
<td>projectteam_employeesatisfaction</td>
<td>241,8630</td>
<td>1753,398</td>
<td>.603</td>
<td>.983</td>
</tr>
<tr>
<td>newemployees_employeesatisfaction</td>
<td>241,9830</td>
<td>1753,269</td>
<td>.608</td>
<td>.983</td>
</tr>
<tr>
<td>projectteam_training</td>
<td>241,7167</td>
<td>1755,384</td>
<td>.575</td>
<td>.983</td>
</tr>
</tbody>
</table>

E5: Test for normal distribution of data

Figure 14: Example bad P-P plot (1)

Figure 15: Example bad P-P plot (2)

Figure 16: Example good P-P plot

Figure 17: Example good histogram
**E6: Correlation matrices**

Because the correlation matrix encompasses 69 variables, it is impossible to provide it in this document. Therefore only some parts of the correlation matrix are shown. For further details see the contact details shown in appendix B.

**Table 16: Part of correlation matrix (1)**

<table>
<thead>
<tr>
<th></th>
<th>Correlations</th>
<th>bigdata_increase_revenue</th>
<th>sales_revenue</th>
<th>retention_revenue</th>
<th>satisfaction_revenue</th>
<th>marketshare_revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>bigdata_increase_revenue</td>
<td>Correlation Coefficient</td>
<td>1.000</td>
<td>0.716</td>
<td>0.633</td>
<td>0.603</td>
<td>0.475</td>
</tr>
<tr>
<td></td>
<td>Sig. (1-tailed)</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>0.716</td>
<td>1.000</td>
<td>0.664</td>
<td>0.595</td>
<td>0.576</td>
</tr>
<tr>
<td>sales_revenue</td>
<td>Sig. (1-tailed)</td>
<td>0.000</td>
<td></td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>0.633</td>
<td>0.664</td>
<td>1.000</td>
<td>0.631</td>
<td>0.355</td>
</tr>
<tr>
<td>retention_revenue</td>
<td>Sig. (1-tailed)</td>
<td>0.000</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>0.603</td>
<td>0.595</td>
<td>0.631</td>
<td>1.000</td>
<td>0.401</td>
</tr>
<tr>
<td>satisfaction_revenue</td>
<td>Sig. (1-tailed)</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>0.475</td>
<td>0.576</td>
<td>0.355</td>
<td>0.401</td>
<td>1.000</td>
</tr>
<tr>
<td>marketshare_revenue</td>
<td>Sig. (1-tailed)</td>
<td>0.006</td>
<td>0.001</td>
<td>0.035</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.05 level (1-tailed).**

**. Correlation is significant at the 0.01 level (1-tailed).**
Table 17: Part of correlation matrix (2)

<table>
<thead>
<tr>
<th>Correlations</th>
<th>crossselling_sales</th>
<th>customized_deals_sales</th>
<th>crossselling_retention</th>
<th>customized_deals_retention</th>
<th>customized_deals_satisfaction</th>
<th>insights_competing_market_marketshare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>1.000</td>
<td>.399*</td>
<td>.563**</td>
<td>.333*</td>
<td>.596**</td>
<td>.571**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.019</td>
<td>.001</td>
<td>.045</td>
<td>.001</td>
<td>.016</td>
<td>.298</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.399*</td>
<td>1.000</td>
<td>.157</td>
<td>.770**</td>
<td>.412*</td>
<td>.107</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.019</td>
<td>.217</td>
<td>.000</td>
<td>.016</td>
<td>.298</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.563**</td>
<td>.157</td>
<td>1.000</td>
<td>.310</td>
<td>.526**</td>
<td>.278</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.001</td>
<td>.217</td>
<td>.000</td>
<td>.090</td>
<td>.080</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.333*</td>
<td>.770**</td>
<td>.310</td>
<td>1.000</td>
<td>.609**</td>
<td>.293</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.045</td>
<td>.000</td>
<td>.058</td>
<td>.000</td>
<td>.069</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.596**</td>
<td>.412*</td>
<td>.526**</td>
<td>.609**</td>
<td>1.000</td>
<td>.431*</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.001</td>
<td>.016</td>
<td>.002</td>
<td>.000</td>
<td>.012</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.571**</td>
<td>.107</td>
<td>.278</td>
<td>.293</td>
<td>.431*</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.001</td>
<td>.298</td>
<td>.080</td>
<td>.069</td>
<td>.012</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

*. Correlation is significant at the 0.05 level (1-tailed).

**. Correlation is significant at the 0.01 level (1-tailed).
### Table 18: Part of correlation matrix (3)

<table>
<thead>
<tr>
<th>Correlations</th>
<th>packages_synchrizingIT</th>
<th>packages_legacy</th>
<th>projectteam_training_errors</th>
<th>newemployees_training_errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>1.000</td>
<td>.782**</td>
<td>.384*</td>
<td>.498**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.</td>
<td>.000</td>
<td>.024</td>
<td>.004</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.782**</td>
<td>1.000</td>
<td>.298</td>
<td>.366*</td>
</tr>
<tr>
<td>packages_legacy</td>
<td>.000</td>
<td>.066</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.384*</td>
<td>.298</td>
<td>1.000</td>
<td>.700**</td>
</tr>
<tr>
<td>projectteam_training_errors</td>
<td>.024</td>
<td>.066</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.498**</td>
<td>.366*</td>
<td>.700**</td>
<td>1.000</td>
</tr>
<tr>
<td>newemployees_training_errors</td>
<td>.004</td>
<td>.030</td>
<td>.000</td>
<td>.</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (1-tailed).

*. Correlation is significant at the 0.05 level (1-tailed).

### Table 19: Part of correlation matrix (4)

<table>
<thead>
<tr>
<th>Correlations</th>
<th>bigdata_decrease_costs</th>
<th>employeecosts_costs</th>
<th>creditscoring_costs</th>
<th>fraud_costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>1.000</td>
<td>.394</td>
<td>.579</td>
<td>.702</td>
</tr>
<tr>
<td>bigdata_decrease_costs</td>
<td>.</td>
<td>.021</td>
<td>.001</td>
<td>.000</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.394</td>
<td>1.000</td>
<td>.728</td>
<td>.289</td>
</tr>
<tr>
<td>employeecosts_costs</td>
<td>.021</td>
<td>.</td>
<td>.000</td>
<td>.072</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.579</td>
<td>.728</td>
<td>1.000</td>
<td>.570</td>
</tr>
<tr>
<td>creditscoring_costs</td>
<td>.001</td>
<td>.000</td>
<td>.</td>
<td>.001</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.702</td>
<td>.289</td>
<td>.570</td>
<td>1.000</td>
</tr>
<tr>
<td>fraud_costs</td>
<td>.000</td>
<td>.072</td>
<td>.001</td>
<td>.</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

*. Correlation is significant at the 0.05 level (1-tailed).

**. Correlation is significant at the 0.01 level (1-tailed).
Table 20: Part of correlation matrix (5)

**Correlations**

<table>
<thead>
<tr>
<th></th>
<th>employeesatisfaction_employeesatisfaction</th>
<th>training_employeesatisfaction</th>
<th>projectteam_employeesatisfaction</th>
<th>newemployees_employeesatisfaction</th>
<th>projectteam_training</th>
</tr>
</thead>
<tbody>
<tr>
<td>employeesatisfaction_employeesatisfaction</td>
<td>Correlation Coefficient: 1.000**</td>
<td>.676**</td>
<td>.649**</td>
<td>.725**</td>
<td>.371**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.000**</td>
<td>.000</td>
<td>.000</td>
<td>.028</td>
<td>.028</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>training_employeesatisfaction</td>
<td>Correlation Coefficient: .676**</td>
<td>1.000**</td>
<td>.456**</td>
<td>.625**</td>
<td>.169**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.000</td>
<td>.008</td>
<td>.000</td>
<td>.200</td>
<td>.200</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>projectteam_employeesatisfaction</td>
<td>Correlation Coefficient: .649</td>
<td>.456</td>
<td>1.000</td>
<td>.618</td>
<td>.778</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.000</td>
<td>.008</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>newemployees_employeesatisfaction</td>
<td>Correlation Coefficient: .725**</td>
<td>.625**</td>
<td>.618</td>
<td>1.000</td>
<td>.354**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.354</td>
<td>.354</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>projectteam_training</td>
<td>Correlation Coefficient: .371</td>
<td>.169</td>
<td>.778</td>
<td>.354</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.028</td>
<td>.200</td>
<td>.000</td>
<td>.035</td>
<td>.035</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (1-tailed).**

*Correlation is significant at the 0.05 level (1-tailed).*

E7: Results of questionnaire

The strength of the relationships between the different variables is shown for in table 21. These results are linked to the conceptual framework shown in figure 5.

Table 21: Results of questionnaire

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase revenue</td>
<td>Increase EBIT</td>
<td>0,793</td>
</tr>
<tr>
<td>Increase sales</td>
<td>Increase revenue</td>
<td>0,719</td>
</tr>
<tr>
<td>Increase retention of customer</td>
<td>Increase revenue</td>
<td>0,741</td>
</tr>
<tr>
<td>Increase customer satisfaction</td>
<td>Increase revenue</td>
<td>0,748</td>
</tr>
<tr>
<td>Increase market share</td>
<td>Increase revenue</td>
<td>0,674</td>
</tr>
<tr>
<td>Increase cross selling</td>
<td>Increase sales</td>
<td>0,726</td>
</tr>
<tr>
<td>Increase cross selling</td>
<td>Increase retention of customer</td>
<td>0,685</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase sales</td>
<td>0,756</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase retention of customer</td>
<td>0,763</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase customer satisfaction</td>
<td>0,762</td>
</tr>
<tr>
<td>Improve insights in competing market</td>
<td>Increase market share</td>
<td>0,726</td>
</tr>
<tr>
<td>Decrease costs</td>
<td>Increase EBIT</td>
<td>0,711</td>
</tr>
<tr>
<td>Decrease employee costs</td>
<td>Decrease costs</td>
<td>0,585</td>
</tr>
<tr>
<td>Decrease wrong credit scoring</td>
<td>Decrease costs</td>
<td>0,704</td>
</tr>
</tbody>
</table>
Table 21: Results of questionnaire (continued)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase fraud detection</td>
<td>Decrease costs</td>
<td>0.807</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Increase cross selling</td>
<td>0.726</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Increase cross selling</td>
<td>0.711</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase cross selling</td>
<td>0.689</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Increase cross selling</td>
<td>0.733</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase cross selling</td>
<td>0.696</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Increase customized deals</td>
<td>0.704</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Increase customized deals</td>
<td>0.719</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase customized deals</td>
<td>0.674</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Increase customized deals</td>
<td>0.719</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase customized deals</td>
<td>0.696</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Improve insights competing market</td>
<td>0.733</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Improve insights competing market</td>
<td>0.704</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Improve insights competing market</td>
<td>0.681</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Improve insights competing market</td>
<td>0.726</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Improve insights competing market</td>
<td>0.689</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Decrease wrong credit scoring</td>
<td>0.748</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Decrease wrong credit scoring</td>
<td>0.726</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Decrease wrong credit scoring</td>
<td>0.704</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Decrease wrong credit scoring</td>
<td>0.726</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Decrease wrong credit scoring</td>
<td>0.674</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Increase fraud detection</td>
<td>0.785</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Increase fraud detection</td>
<td>0.719</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase fraud detection</td>
<td>0.689</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Increase fraud detection</td>
<td>0.733</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase fraud detection</td>
<td>0.667</td>
</tr>
<tr>
<td>Provide training in business changes</td>
<td>Decrease employee costs</td>
<td>0.656</td>
</tr>
<tr>
<td>Increase employee satisfaction</td>
<td>Decrease employee costs</td>
<td>0.600</td>
</tr>
<tr>
<td>Attract new employees</td>
<td>Increase employee satisfaction</td>
<td>0.616</td>
</tr>
<tr>
<td>Create project team</td>
<td>Increase employee satisfaction</td>
<td>0.640</td>
</tr>
<tr>
<td>Create project team</td>
<td>Provide training in business changes</td>
<td>0.669</td>
</tr>
<tr>
<td>Attract new employees</td>
<td>Provide training in use new IT</td>
<td>0.685</td>
</tr>
<tr>
<td>Create project team</td>
<td>Provide training in use new IT</td>
<td>0.700</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about payments</td>
<td>0.644</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about savings</td>
<td>0.644</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about financings</td>
<td>0.652</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about investments</td>
<td>0.654</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about insurances</td>
<td>0.644</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about payments</td>
<td>0.733</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about savings</td>
<td>0.719</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about financings</td>
<td>0.733</td>
</tr>
</tbody>
</table>
Table 21: Results of questionnaire (continued)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchronize IT</td>
<td>Information about investments</td>
<td>0.726</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about insurances</td>
<td>0.726</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about savings</td>
<td>0.719</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about financings</td>
<td>0.733</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about investments</td>
<td>0.726</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about insurances</td>
<td>0.726</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about payments</td>
<td>0.778</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about savings</td>
<td>0.770</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about financings</td>
<td>0.785</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about investments</td>
<td>0.793</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about insurances</td>
<td>0.785</td>
</tr>
<tr>
<td>Installing Big Data packages</td>
<td>Synchronize IT</td>
<td>0.692</td>
</tr>
<tr>
<td>Installing Big Data packages</td>
<td>Combine legacy and new hardware and software</td>
<td>0.669</td>
</tr>
</tbody>
</table>

Appendix F: Big Data Value Framework for the Financial Services Industry
This appendix shows the results after the adaptations made to the conceptual framework. Table 22 shows all the weights of the lines that are present in the framework. The Big Data Value Framework for the Financial Services Industry is shown in figure 9.

Table 22: Weights of lines after adaptation

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase revenue</td>
<td>Increase EBIT</td>
<td>0.793</td>
</tr>
<tr>
<td>Increase sales</td>
<td>Increase revenue</td>
<td>0.719</td>
</tr>
<tr>
<td>Increase retention of customer</td>
<td>Increase revenue</td>
<td>0.741</td>
</tr>
<tr>
<td>Increase customer satisfaction</td>
<td>Increase revenue</td>
<td>0.748</td>
</tr>
<tr>
<td>Increase market share</td>
<td>Increase revenue</td>
<td>0.674</td>
</tr>
<tr>
<td>Increase cross selling</td>
<td>Increase sales</td>
<td>0.726</td>
</tr>
<tr>
<td>Increase cross selling</td>
<td>Increase retention of customer</td>
<td>0.685</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase sales</td>
<td>0.756</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase retention of customer</td>
<td>0.763</td>
</tr>
<tr>
<td>Increase customized deals</td>
<td>Increase customer satisfaction</td>
<td>0.762</td>
</tr>
<tr>
<td>Improve insights in competing market</td>
<td>Increase market share</td>
<td>0.726</td>
</tr>
<tr>
<td>Decrease costs</td>
<td>Increase EBIT</td>
<td>0.711</td>
</tr>
<tr>
<td>Decrease wrong credit scoring</td>
<td>Decrease costs</td>
<td>0.704</td>
</tr>
<tr>
<td>Increase fraud detection</td>
<td>Decrease costs</td>
<td>0.807</td>
</tr>
</tbody>
</table>
Table 22: Weights of lines after adaptation (continued)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information about payments</td>
<td>Increase cross selling</td>
<td>0,726</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Increase cross selling</td>
<td>0,711</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase cross selling</td>
<td>0,689</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Increase cross selling</td>
<td>0,733</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase cross selling</td>
<td>0,696</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Increase customized deals</td>
<td>0,704</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Increase customized deals</td>
<td>0,719</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase customized deals</td>
<td>0,674</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Increase customized deals</td>
<td>0,719</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase customized deals</td>
<td>0,696</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Improve insights competing market</td>
<td>0,733</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Improve insights competing market</td>
<td>0,704</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Improve insights competing market</td>
<td>0,681</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Improve insights competing market</td>
<td>0,726</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Improve insights competing market</td>
<td>0,689</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Decrease wrong credit scoring</td>
<td>0,748</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Decrease wrong credit scoring</td>
<td>0,726</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Decrease wrong credit scoring</td>
<td>0,704</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Decrease wrong credit scoring</td>
<td>0,726</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Decrease wrong credit scoring</td>
<td>0,674</td>
</tr>
<tr>
<td>Information about payments</td>
<td>Increase fraud detection</td>
<td>0,785</td>
</tr>
<tr>
<td>Information about savings</td>
<td>Increase fraud detection</td>
<td>0,719</td>
</tr>
<tr>
<td>Information about financings</td>
<td>Increase fraud detection</td>
<td>0,689</td>
</tr>
<tr>
<td>Information about investments</td>
<td>Increase fraud detection</td>
<td>0,733</td>
</tr>
<tr>
<td>Information about insurances</td>
<td>Increase fraud detection</td>
<td>0,667</td>
</tr>
<tr>
<td>Provide training in business changes</td>
<td>Decrease costs</td>
<td>0,656</td>
</tr>
<tr>
<td>Create project team</td>
<td>Provide training in business changes</td>
<td>0,669</td>
</tr>
<tr>
<td>Attract new employees</td>
<td>Provide training in use new IT</td>
<td>0,685</td>
</tr>
<tr>
<td>Create project team</td>
<td>Provide training in use new IT</td>
<td>0,700</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about payments</td>
<td>0,644</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about savings</td>
<td>0,644</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about financings</td>
<td>0,652</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about investments</td>
<td>0,654</td>
</tr>
<tr>
<td>Provide training in use new IT</td>
<td>Information about insurances</td>
<td>0,644</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about payments</td>
<td>0,733</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about savings</td>
<td>0,719</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about financings</td>
<td>0,733</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about investments</td>
<td>0,726</td>
</tr>
<tr>
<td>Synchronize IT</td>
<td>Information about insurances</td>
<td>0,726</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about payments</td>
<td>0,726</td>
</tr>
</tbody>
</table>
Table 22: Weights of lines after adaptation (continued)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about savings</td>
<td>0.719</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about financings</td>
<td>0.733</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about investments</td>
<td>0.726</td>
</tr>
<tr>
<td>Combine legacy and new hardware and software</td>
<td>Information about insurances</td>
<td>0.726</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about payments</td>
<td>0.778</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about savings</td>
<td>0.770</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about financings</td>
<td>0.785</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about investments</td>
<td>0.793</td>
</tr>
<tr>
<td>Visualization and analytics</td>
<td>Information about insurances</td>
<td>0.785</td>
</tr>
<tr>
<td>Installing Big Data packages</td>
<td>Synchronize IT</td>
<td>0.692</td>
</tr>
<tr>
<td>Installing Big Data packages</td>
<td>Combine legacy and new hardware and software</td>
<td>0.669</td>
</tr>
</tbody>
</table>